```
 __   __  _____   _____   _____   _____   __  __
|  \ /  ||  ___ \ |  ____| |  ___ \ |  ____| |  |/  |
|   V   || |___) || |____  | |___) || |      |     /
| |\ /| ||  ____/ |____  | |  __  / | |      |     \
| | V | || |       ____| | | |  \ \ | |____  |  |\  |
|_|   |_||_|      |_____| |_|   |_||_____| |__| \_|  (TM)
```

****

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

****

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 21 15:40:20 2000;   MasPar time 25.93 Seconds
                                            936.964 Million cell updates/sec

Tabular output not generated.

Title: >US-09-235-416-1
Description: (1-784) from US09235416A.pep
Perfect Score: 5422
Sequence: 1 MSGGGNIKVVVRVRPFNARE.........ELRQQQAQMEEALKTAKQEF 784

Scoring table: PAM 150
               Gap 11

Searched: 85561 seqs, 30989116 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: swiss-prot38
               1:swissprot

Statistics: Mean 54.902;   Variance 108.067;   scale 0.508

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2217	40.9	1150	1	KF1B_MOUSE	KINESIN-LIKE PROTEIN K	0.00e+00
2	2175	40.1	1690	1	KF1A_HUMAN	KINESIN-LIKE PROTEIN K	0.00e+00
3	2176	40.1	1695	1	KF1A_MOUSE	KINESIN-LIKE PROTEIN K	0.00e+00
4	2051	37.8	1103	1	KF1C_HUMAN	KINESIN-LIKE PROTEIN K	0.00e+00
5	1975	36.4	1584	1	U104_CAEEL	KINESIN-LIKE PROTEIN U	0.00e+00
6	1008	18.6	699	1	KI22_STRPU	KINESIN-II 85 KDA SUBU	3.40e-173
7	962	17.7	701	1	KF3A_MOUSE	KINESIN-LIKE PROTEIN K	1.37e-163
8	958	17.7	702	1	KF3A_HUMAN	KINESIN-LIKE PROTEIN K	9.35e-163
9	962	17.7	742	1	KI21_STRPU	KINESIN-II 95 KDA SUBU	1.37e-163
10	953	17.6	955	1	KINL_LEICH	KINESIN-LIKE PROTEIN K	1.03e-161
11	945	17.4	747	1	KF3B_HUMAN	KINESIN-LIKE PROTEIN K	4.78e-160
12	942	17.4	747	1	KF3B_MOUSE	KINESIN-LIKE PROTEIN K	2.02e-159
13	897	16.5	784	1	KL68_DROME	KINESIN-LIKE PROTEIN F	1.60e-150
14	870	16.0	786	1	FL10_CHLRE	KINESIN-LIKE PROTEIN F	1.84e-144
15	846	15.6	1231	1	KIF4_MOUSE	KINESIN-LIKE PROTEIN K	1.71e-139
16	839	15.5	672	1	OSM3_CAEEL	KINESIN-LIKE PROTEIN O	4.80e-138
17	842	15.5	1232	1	KIF4_HUMAN	KINESIN-LIKE PROTEIN K	1.15e-138
18	815	15.0	928	1	KINH_NEUCR	KINESIN HEAVY CHAIN.	4.35e-133
19	799	14.7	2663	1	CENE_HUMAN	CENTROMERIC PROTEIN E	8.66e-130
20	762	14.1	805	1	YGW6_YEAST	PUTATIVE KINESIN-LIKE	3.54e-122
21	766	14.1	935	1	KINH_SYNRA	KINESIN HEAVY CHAIN (S	5.34e-123
22	754	13.9	883	1	YB3D_SCHPO	PUTATIVE KINESIN-LIKE	1.55e-120
23	747	13.8	160	1	KF1C_MOUSE	KINESIN-LIKE PROTEIN K	4.24e-119
24	741	13.7	967	1	KINH_LOLPE	KINESIN HEAVY CHAIN.	7.19e-118
25	731	13.5	1031	1	KINH_STRPU	KINESIN HEAVY CHAIN.	8.04e-116
26	727	13.4	793	1	KF3C_HUMAN	KINESIN-LIKE PROTEIN K	5.29e-115
27	724	13.4	796	1	KF3C_MOUSE	KINESIN-LIKE PROTEIN K	2.18e-114
28	725	13.4	815	1	KINH_CAEEL	KINESIN HEAVY CHAIN.	1.36e-114
29	715	13.2	796	1	KF3C_RAT	KINESIN-LIKE PROTEIN K	1.51e-112
30	707	13.0	975	1	KINH_DROME	KINESIN HEAVY CHAIN.	6.48e-111
31	704	13.0	1032	1	KINH_HUMAN	NEURONAL KINESIN HEAVY	2.66e-110
32	697	12.9	1027	1	KINN_MOUSE	NEURONAL KINESIN HEAVY	7.12e-109
33	665	12.3	963	1	KINH_HUMAN	KINESIN HEAVY CHAIN (U	2.32e-102
34	665	12.3	963	1	KINH_MOUSE	KINESIN HEAVY CHAIN (U	2.32e-102
35	654	12.1	1057	1	EG5_HUMAN	KINESIN-RELATED MOTOR	3.97e-100
36	656	12.1	1060	1	EG5_XENLA	KINESIN-RELATED MOTOR	1.56e-100
37	648	12.0	776	1	KLP1_CHLRE	KINESIN-LIKE PROTEIN K	6.53e-99
38	653	12.0	1066	1	KL61_DROME	BIPOLAR KINESIN KRP-13	6.33e-100
39	652	12.0	1067	1	EG52_XENLA	KINESIN-LIKE PROTEIN E	1.01e-99
40	649	12.0	1184	1	BIMC_EMENI	KINESIN-LIKE PROTEIN B	4.09e-99
41	645	11.9	643	1	CTK2_XENLA	CARBOXY-TERMINAL KINES	2.65e-98
42	627	11.6	754	1	KATC_ARATH	KINESIN-LIKE PROTEIN C	1.16e-94
43	623	11.5	745	1	KATB_ARATH	KINESIN-LIKE PROTEIN B	7.46e-94
44	616	11.4	770	1	KLPA_EMENI	KINESIN-LIKE PROTEIN K	1.93e-92
45	602	11.1	832	1	KLP1_SCHPO	KINESIN-LIKE PROTEIN 1	1.28e-89

ALIGNMENTS

RESULT   1
ID   KF1B_MOUSE   STANDARD;   PRT;   1150 AA.
AC   Q60575;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF1B.
GN   KIF1B.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95094296.
RA   Nangaku M., Sato-Yoshitake R., Okada Y., Noda Y., Takemura R.,
RA   Yamazaki H., Hirokawa N.;
RT   "KIF1B, a novel microtubule plus end-directed monomeric motor protein
RT   for transport of mitochondria.";
RL   Cell 79:1209-1220(1994).
CC   -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC   MICROTUBULE PLUS END-DIRECTED MOTILITY.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND
CC   TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM.
CC   -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
CC   TISSUE (MAINLY IN THE CEREBELLUM AND CEREBRUM) WITHIN A SINGLE
CC   TYPE OF NEURONAL CELL.
CC   -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC   SUBFAMILY.

DR   EMBL; D17577; BAA04503.1; -.
DR   HSSP; P56536; 2KIN.
DR   MGD; MGI:108426; Kif1b.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
```

**Left column:**

```
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR   PROSITE; PS50006; FHA_DOMAIN; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
FT   DOMAIN      1    355    MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN    359    380    COILED COIL (POTENTIAL).
FT   DOMAIN    424    456    COILED COIL (POTENTIAL).
FT   DOMAIN    510    566    COILED COIL (POTENTIAL).
FT   DOMAIN    622    663    FHA.
FT   DOMAIN    814    858    COILED COIL (POTENTIAL).
FT   NP_BIND    97    104    ATP (POTENTIAL).
SQ   SEQUENCE  1150 AA;  130278 MW;  5AA426DAFEBD252F CRC64;

Query Match             40.9%;  Score  2217;  DB 1;  Length 1150;
Best Local Similarity   55.3%;  Pred. No. 0.00e+00;
Matches  340;  Conservative 132;  Mismatches 114;  Indels  29;  Gaps  20;

Db      1  MSGAS-VKVAVRVRPFNSRETSKESKCIIQMOGNSTSIINP----K--NP-K---EAPK   48
              :|||: :|| ||||||:||||||||:|: :|:| : ::||
Qy      1  MSGGGNIKVVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK   60

Db     49  SFSFDYSYWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNVCIFAYGQTGAGKSYTMM  108
           :|:|| |||| :|:  :||||||||||||||||||||||||||||||||||||
Qy     61  AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db    109  GKQEESQAVIIPQLCEELFEKIND-NCNEEMSYSVEVSYMEIYCERVRDLLNPKNKGNLR  167
           |:||  |:|||| :||: ||| ::||||:|| ||||||||||||||||||| ||||
Qy    120  GYGKE-HGVI-PRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK  177

Db    168  VREHPLLGPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQ  227
           ||||||||||||||||||||| :| :|||||:||||||||||||||||||||||||: :|:
Qy    178  VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ  237

Db    228  KKQDPETNLSTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE  287
           ||  :|:|| |||| ||:|||||||||||||:||:: ||||::| ||:||||:|||:||
Qy    238  KRHDAETNITTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE  297

Db    288  VSK-KKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRYADRAK  346
           |:| ||| |||||||||||||||:|||||||||| |::|||||| |:|||||||||:|
Qy    298  MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK  357

Db    347  QIKCNAVINEDPNAKLVRELKEEVTRLKDLLRAQGLGDI-IDTSMGSLTSS-PSSCCSLNS  404
           |:| ||:| ||| ||||||:|:: :| :: :|: :|
Qy    358  RIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGSGGPVEESYPPDTPLEK         417

Db    405  QVGLTSVTSIQPRIMSTPGGEEAEIERLKESEKIIAELNETWEEKLRKTEAIRMEREALLA  464
           |:|: :|| :|| ||: :||:|:|:|:| ||||||: |: ||:||||||:
Qy    418  QI-V-SIQQPDATVKKMSKAE-IVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALE  474

Db    465  EMGVAIREDRGDIGVFSPKKTPHLVNLNEDPLMSECLLYYIKDGITRVGQADAERRQDIV  524
           |:| |: : :|||: ||| : | :| :| ||| : ||
Qy    475  ELGISI-E-KGFVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNVQDTQAEIR  532

Db    525  LSGAHIKEEHCLFRSERSNTGEVIVLEPCERSETYVNGKRVAHPVQLRSGNRIIMGKNH  584
           ||| |||| : :|: :  :||::: |: :|||||:|:|||:| |
Qy    533  LNGSKILKEHCTF--E--NVDNV-VTVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFH  587

Db    585  VFRFNHPEQARARE                                              599
           :||||||:
Qy    588  IFRFNHPEEAARAERQ                                            602

RESULT  2
ID   KF1A_HUMAN      STANDARD;      PRT;   1690 AA.
AC   Q12756;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF1A (AXONAL TRANSPORTER OF SYNAPTIC VESICLES).
GN   KIF1A OR ATSV.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
```

Right column:

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 96299637.
RA   Furlong R.A., Zhou C.Y., Ferguson-Smith M.A., Affara N.A.;
RT   "Characterization of a kinesin-related gene ATSV, within the tuberous
RT   sclerosis locus (TSC1) candidate region on chromosome 9Q34.";
RL   Genomics 33:42-429(1996).
CC   -!- FUNCTION: MOTOR FOR ANTEROGRADE AXONAL TRANSPORT OF SYNAPTIC
CC       VESICLE PRECURSORS (BY SIMILARITY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC       SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; X90840; CAA62346.1; -.
DR   HSSP; P17119; 3KAR.
DR   MIM; 601255; -.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00169; PH; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
DR   PROSITE; PS50006; FHA_DOMAIN; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
FT   DOMAIN      1    361    MECHANOCHEMICAL (MOTOR).
FT   DOMAIN    366    383    COILED COIL (POTENTIAL).
FT   DOMAIN    429    462    COILED COIL (POTENTIAL).
FT   DOMAIN    516    572    FHA.
FT   DOMAIN    622    681    COILED COIL (POTENTIAL).
FT   DOMAIN    801    822    PH.
FT   DOMAIN   1575   1673    FHA_DOMAIN; 1.
FT   NP_BIND    97    104    ATP (POTENTIAL).
SQ   SEQUENCE  1690 AA;  191083 MW;  D8DDEC784624FB4D CRC64;

Query Match             40.1%;  Score  2175;  DB 1;  Length 1690;
Best Local Similarity   54.1%;  Pred. No. 0.00e+00;
Matches  335;  Conservative 131;  Mismatches 122;  Indels  31;  Gaps  19;

Db      1  MAGAS-VKVAVRVRPFNSREMSRDSKCIIQMSGSTTIVNP----K-Q-P-K---ETPK   48
           |:|||: :|| :|||||||:||||||||:|:::|:| :|| :||
Qy      1  MSGGGNIKVVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK   60

Db     49  SFSFDYSYWSHTSPEDINTASQKQYRDIGEEMLQHAFEGYNVCIFAYGQTGAGKSYTMM  108
           :|:|| |||| :|: :|| :|:||:||:|||||||||||||||||||||||||
Qy     61  AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db    109  GKQEKDQQGIIPQLCEDLFSRINDTTND-NMSYSVEVSYMEIYCERVRDLLNPKNKGNLR  167
           | :|: ::|:|| :|| ||: ::||:|| ||||||||||||||||||||| ||||
Qy    120  G-YGKEH-GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK  177

Db    168  VREHPLLGPYVEDLSKLAVTSYNDIQDLMDSGNKARTVAATNMNETSSRSHAVFNIIFTQ  227
           ||||| |||||||:||:|| :||:||:||||||||||||||||||||||||||: :||:
Qy    178  VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ  237

Db    228  KRHDAETNITTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE  287
           | ::|:|:|||:|||::|||||||||||:|||| |:||||||::| ||:||:||:|||||
Qy    238  KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD  297

Db    288  MDSGPKNKKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRY  347
           |:||  :|:|: :|:|||||||||||:|:||||| ||:|| ||||||: |:|||||||
Qy    298  MSSG---KQKKN-QLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRY  352
```

```
Db  348 ADRAKQIRCNAVINEDPNNKLIRELKDEVTRLRD-LLYAQGLGDITDMTNALVGMSPSSS 406
        ||:|:: :||:|||| :|||||:::|| | : |:: | : : |:|
Qy  353 ADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPD 412

Db  407 LSALSSRAASVSSLHERILFAPGSEEAIERLKETEKIIAELNETWEEKLRRTEAIRMERE 466
        ||:|| : |: |||||:||||| | |::||||| | |::||
Qy  413 -TPLEKQIVSIQQPDATVK-KMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKERE 470

Db  467 ALLAEMGVAMREDGGTLGVFSPKKTPHLVNLNEDPLMSECLLYYIKDGITRVGREDGERR 526
        | | : | : | ||||||:|:|| || | :: :
Qy  471 AALEELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQ 528

Db  527 QDIVLSGHFIKEEHCVFRSDSRGGSEAVVTLEPCEGADTYVNGKKVTEPSILRSGNRIIM 586
        ::| |: |: ||||| | :| :| : |:|| | : |||| : |||
Qy  529 AEIRLNGSKILKEHCTF--EN--VDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIIL 583

Db  587 GKSHVFRFTHPEQARQERE 605
        | :|||| ||:|:||:|| :||
Qy  584 GDFHIFRFNHPEEARAERQ 602
```

```
DR   MGD; MGI:108391; KIF1A.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00169; PH; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
DR   PROSITE; PS50006; FHA_DOMAIN; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
FT   DOMAIN       1    361       MECHANOCHEMICAL (MOTOR).
FT   DOMAIN     366    383       COILED COIL (POTENTIAL).
FT   DOMAIN     429    462       COILED COIL (POTENTIAL).
FT   DOMAIN     516    572       FHA.
FT   DOMAIN     622    681       COILED COIL (POTENTIAL).
FT   DOMAIN     801    822       COILED COIL (POTENTIAL).
FT   DOMAIN    1580   1678       PH.
FT   NP_BIND     97    104       ATP (BY SIMILARITY)
SQ   SEQUENCE   1695 AA;  191724 MW;  D6EC3B88CBC9CCC6 CRC64;
```

```
Query Match             40.1%;  Score 2176;  DB 1;  Length 1695;
Best Local Similarity   54.0%;  Pred. No. 0.00e+00;
Matches  334;  Conservative  133;  Mismatches  121;  Indels  31;  Gaps  19;
```

```
Db    1 MAGAS-VKVAVRVRPFNSREMSRDSKCIIQMSGSTTTIVNP----K-Q-P-K---ETPK 48
         |:|:: :|| ||||||||| :|||:: | :|||:|  |: :  |  :|  :    ||
Qy    1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60

Db   49 SFSFDYSYWSHTSPEDINYASQKQYRDIGEEMLQHAFEGYNVCIFAYGQTGAGKSYTMM 108
         :|:|| |||| : ||  :| : : :|:|:  ||| |::||: ||||||||||||| :|
Qy   61 AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db  109 GKQEKDQQGIIPQLCEDLFSRINDTTND-NMSYSVEVSYMEIYCERVRDLLNPKNKGNLR 167
         | | : | || : |||| | :: ||:| :|::|  ::|||:|:| ||||||| : ||||
Qy  120 G-YGKEH-GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK 177

Db  168 VREHPLLGPYVEDISKLLAVTSYNDIQDLMDSGNKPRTVAATNMNETSSRSHAVFNIIFTQ 227
         |||| ||||||||:|| ||| :|||:| |||:|:| || ||||||||||||||||| : |:
Qy  178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237

Db  228 KRHDAETNITTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE 287
         :|: ||| : | :| ::|||||||||||| |: |: ||:|| :|||| | |::||||
Qy  238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db  288 MDSGPNKNKKKKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRY 347
         | | |:  :  :  : :|||||||||| ||||||||:|||:||:||||| ::|||||||
Qy  298 MSSG----KQKKN-QLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRY 352

Db  348 ADRAKQIRCNAVINEDPNNKLIRELKDEVTRLRD-LLYAQGLGDITDMTNALVGMSPSSS 406
         ||:|:: :||:|||| :|||||:::|| | : |:: | : : |:|
Qy  353 ADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPD 412

Db  407 LSALSSRAASVSSLHERILFAPGSEEAIERLKETEKIIAELNETWEEKLRRTEAIRMERE 466
         ||:|| : |: |||||:||||| | |::||||| | |::||
Qy  413 -TPLEKQIVSIQQPDATVK-KMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKERE 470

Db  467 ALLAEMGVAMREDGGTLGVFSPKKTPHLVNLNEDPLMSECLLYYIKDGITRVGREDAERR 526
         | | : | : | ||||||:|:|| || | :: :
Qy  471 AALEELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQ 528

Db  527 QDIVLSGHFIKEEHCIFRSDSRGGEAVVTLEPCEGADTYVNGKKVTEPSILRSGNRIIM 586
         ::| |: |: |||||  | :| :| : |:|| | : |||| : |||
Qy  529 AEIRLNGSKILKEHCTF--EN--VDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIIL 583

Db  587 GKSHVFRFNHPEQARQERE 605
         | :|||| ||:|:||:||
Qy  584 GDFHIFRFNHPEEARAERQ 602
```

```
RESULT 3
ID   KIF1A_MOUSE    STANDARD;      PRT;   1695 AA.
AC   P33173; Q61770;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF1A.
GN   KIF1A OR KIF1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 95292344.
RA   Okada Y., Yamazaki H., Sekine-Aizawa Y., Hirokawa N.;
RT   "The neuron-specific kinesin superfamily protein KIF1A is a unique
RT   monomeric motor for anterograde axonal transport of synaptic vesicle
RT   precursors.";
RL   Cell 81:769-780(1995).
RN   [2]
RP   PRELIMINARY SEQUENCE OF 100-247 FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93077686.
RA   Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M.,
RA   Hirokawa N.;
RT   "Kinesin family in murine central nervous system.";
RL   J. Cell Biol. 119:1287-1296(1992).
CC   -!- FUNCTION: MOTOR FOR ANTEROGRADE AXONAL TRANSPORT OF SYNAPTIC
CC       VESICLE PRECURSORS.";
CC   -!- SUBUNIT: MONOMER.
CC   -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
CC       TISSUE (MAINLY IN THE CEREBELLUM AND CEREBRUM) WITHIN A SINGLE
CC       TYPE OF NEURONAL CELL. WITHIN THE NEURONAL CELL LEVELS ARE
CC       CONCENTRATED AROUND THE AXON, WITH SMALLER AMOUNTS IN THE
CC       PERINCLEAR AND SYNAPTIC REGIONS.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC       SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; D29951; BAA06221.1; -.
DR   PIR; E44259; E44259.
DR   HSSP; P17119; 3KAR.

RESULT 4
ID   KIF1C_HUMAN    STANDARD;      PRT;   1103 AA.
AC   O43896;
DT   15-DEC-1998 (Rel. 37, Created)
```

```
Db   228  RCHDQLTGLDSEKVSKISLVDLAGSERADSSGARGMRLKEGANINKSLTTLGKVISALAD  287
          : ||:  |:|||:|||||||||||||||| ::||  |: |||||||  |:|| ||:|| ||
Qy   238  KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD  297

Db   288  MQS-KKRKSDFIPYRDSVLTWLLKENLGGNSRTAMIAALSPADINYEETLSTLRYADRTK  346
          | : ||: | :|||||||||||||::|||||:||||||:||||||:||||||||||||:|
Qy   298  MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK  357

Db   347  QIRCNAIINEDPNARLIREIQEEVARLRELLMAQGLSASALEGLKTEEGSVRGALPAVSS  406
          :|: : :|:|||||:|||:::|::|:|: |                             :
Qy   358  RIKNHAVVNEDPNARMIRELKEELAQLRSKL--QS-SGGG-GG--GAGSG-GPVEE-SY   409

Db   407  PPAPVSPSSPTTHNGELEPSFSPNTESQIGPEEAMERLQETEKIIAELNETWEEKLRKTE  466
          || :  :    ::  : |::  :  :::: ::| |::::| :|::|:||:|||||| ||
Qy   410  PPD--TPLE--KQIVSIQQP-DA-TVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTE  463

Db   467  ALRMEREALLABMGVAVREDGGTVGVFSEPKKTPHLVNLNEDPLMSECLLYHIKDGVTRVG  526
          ::::||||:||: |:: ||| |  :  : ::|||:|||::||||:||||:|:||| |||||
Qy   464  EIHKEREAALEELGISI-EKG-FVGPYHSKEMPHIVNLSDDPLLAECLVYNIKPGQTRVG   521

Db   527  QV--DM--DIKLTGQFIREQHCLFRSIPQPDGEVVVTLEPCEGAETYVNGKLVTEPLVLK  582
          :| : |  |:||:| ||:|::|::|::   : ||| |  |  |:    ::|:|:|:||||
Qy   522  NVNQDTQAEIRLNGSKILKEHCTFENV---D-NVV-TIVPNEKAAVMVNGVRIDKPTRLR  576

Db   583  SGNRIVMGKNHVFRFNHPEQARLERE  608
          || ||::|.::|:|||  |::| | :
Qy   577  SGYRIILGDFHIFRFNPNHPEEARAERQ  602
```

```
RESULT   5
ID   U104_CAEEL     STANDARD;      PRT;   1584 AA.
AC   P23678;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   KINESIN-LIKE PROTEIN UNC-104.
GN   UNC-104.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91097805.
RA   Otsuka A.J., Jeyaprakash A., Garcia-Anoveros J., Tang L.Z., Fisk G.,
RA   Hartshorne T., Franco R., Born T.;
RT   "The C. elegans unc-104 gene encodes a putative kinesin heavy
RT   chain-like protein.";
RL   Neuron 6:113-122(1991).
CC   -!- FUNCTION: INVOLVED IN MICROTUBULE-ASSOCIATED TRANSPORT.
CC   -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M58582; AAA03517.1; -.
DR   PIR; JN0114; JN0114.
DR   HSSP; P55536; 2KIN.
DR   PFAM; PF00169; PH; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW   Transport.
FT   DOMAIN        1    354       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY)..
```

```
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF1C.
GN   KIF1C.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HIPPOCAMPUS;
RX   MEDLINE; 98352063.
RA   Dorner C., Ciossek T., Mueller S., Moeller N.P.H., Ullrich A.,
RA   Lammers R.;
RT   "Characterization of KIF1C, a new kinesin-like protein involved in
RT   vesicle transport from the Golgi apparatus to the endoplasmic
RT   reticulum.";
RL   J. Biol. Chem. 273:20267-20275(1998).
CC   -!- FUNCTION: MOTOR REQUIRED FOR THE RETROGRADE TRANSPORT OF GOLGI
CC       VESICLES TO THE ENDOPLASMIC RETICULUM. HAS A MICROTUBULE PLUS END-
CC       DIRECTED MOTILITY.
CC   -!- SUBUNIT: MONOMER (POTENTIAL).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED, WITH MOST
CC       ABUNDANT EXPRESSION IN HEART AND SKELETAL MUSCLE.
CC   -!- PTM: PHOSPHORYLATED ON TYROSINE.
CC   -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC       SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U91329; AAC52117.1; -.
DR   MIM; 603060; -.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil;
KW   Phosphorylation.
FT   DOMAIN        1    355       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN      359    388       COILED COIL (POTENTIAL).
FT   DOMAIN      438    479       COILED COIL (POTENTIAL).
FT   DOMAIN      523    575       FHA.
FT   DOMAIN      633    674       COILED COIL (POTENTIAL).
FT   DOMAIN      828    872       COILED COIL (POTENTIAL).
FT   NP_BIND      97    104       ATP (POTENTIAL).
SQ   SEQUENCE   1103 AA;  123071 MW;  F14BC398D4B2EC10 CRC64;
```

```
Query Match             37.8%;  Score 2051;  DB 1;  Length 1103;
Best Local Similarity   54.3%;
Matches  340;  Conservative 127;  Mismatches 117;  Indels 42;  Gaps 28;

Db     1  MAGAS-VKVAVRVRPFNARETSQDAKCVVSMQGNTTSIINP----K-Q-S-K--DAPK  48
          | :::  ||:||||||||||||| :|:||| ||: ::|:   :       : |   |||
Qy     1  MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK  60

Db    49  SFTFDYSYWSHTSTEDPQFASQQQVRDIGEEMLLHAFEGYNVCIFAYGQTGAGKSYTMM  108
          :|:||| |||   |: ||: :::|:|| | ::|| ||.||| ||||||||||:|||:||
Qy    61  AFAFDRSYWSF-DKNAPNYAREDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db   109  GRQEPGQQGIVPQLCEDLFSRVSENQQSAQ-LSYSVEVSYMEIYCERVRDLLNPKSRGSLR  167
          | :    :| :|::|:|:| ::::: ||  ||:|::::::::||| ||:|| |||:::| ||::|
Qy   120  G-YGK-EHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK  177

Db   168  VREHPILGPYVQDLSKLAVTSVADIADLMDCGNKARTVAATNMNETSSRSHAVFTIVFTQ  227
          |||||  ||||:||:||.|.| ||:|:||| ||||||||||||||||||||||||||:||:|
Qy   178  VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ  237
```

```
FT   DOMAIN      425    445      COILED COIL (POTENTIAL).
FT   DOMAIN      598    652      COILED COIL (POTENTIAL).
FT   DOMAIN      777    797      COILED COIL (POTENTIAL).
FT   DOMAIN      183    335      MICROTUBULE-BINDING.
FT   DOMAIN     1460   1558      PH.
FT   DOMAIN      957   1052      ARG/LYS-RICH (BASIC).
FT   DOMAIN     1203   1584      ARG/LYS-RICH (BASIC).
FT   NP_BIND      93    100      ATP (BY SIMILARITY).
FT   VARIANT     598    598      I -> T.
FT   VARIANT     930    930      V -> M.
SQ   SEQUENCE  1584 AA;  179736 MW;  559462B3FD029B43 CRC64;
```

```
Query Match       36.4%;    Score 1975;   DB 1;   Length 1584;
Best Local Similarity   52.6%;    Pred. No. 0.00e+00;
Matches     318;   Conservative   123;   Mismatches 126;   Indels   38;   Gaps   20;

Db       2  SSVKVAVRVRPFNQREISNTSKCVLQVNGNTTTIN--G--HSINKE-N-F-----SFNF  49
               ::|| ||||||||| |||||||:|:||| | : :   |  : :||
Qy       5  GNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAFAF  64

Db      50  DHSTWSFARNDPHFITQKQVYEELGVEMLEHAFEGYNVCIFAYGQTGSGKSYTMMGKAND  109
               |::||| :|| : | | ||||||||| | ||:|||| |||||||||||||||||:|||
Qy      65  DRSTWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY--YG  122

Db     110  PDEMGIIPRLCNDLFARIDN-NNDKDVQYSVEVSMEIYCERVKDLLNPNSGGNLRVREH  168
               : |||||| : :: .|: : ||:||||||||:|||: ||||||||||||||:||||
Qy     123  KEH-GVIPRICQDDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH  181

Db     169  PLLGPYVDDLTKMAVCSYHDICNLMDEGNKARTVAATNMNSTSSRSHAVFTIVLTQKRHC  228
               ||| || :||:| :|:: :  |||||||||||||||:|:|:||:||||||:|||:||
Qy     182  PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD  241

Db     229  ADSNLDTEKHSKISLVDLAGSERANSTGAEQORLKEGANINKSLTTLGLVISKLAEESTK  288
               : |||:| | |||||||||||||| || || |||||||:|||||||||| :||||:||
Qy     242  EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG  301

Db     289  KKKSNKGVIPYRDSVLTWLLRENLGGNSKTAMLAALSPADINFDETLSTLRYADRAKQIV  348
               |:| :|| || ||||||||||||||:|||||||||||||||||||||||||| | |:|
Qy     302  KQKKNQ-LVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK  360

Db     349  CQAVVNEDPNAKLIRELNEEVIKLIKHHILKDKG-----IDVTD--VQET--PGK--HKK-G  396
               |:| |||||||| |||| | :|| ||:|     |   :  |     : | | :  ||
Qy     361  NHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGGSGGGGGGGSGGGGGGGSGGGGGGGGS  420

Db     397  P-KLP-AHVH----EQL-EKLQESEKLMAEIGKTWEQKLIHTEIRKQREEELRDMGLAC  449
               | : |          |: : :||||| :||:| | ||||| :
Qy     421  SIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALEELGIS-  479

Db     450  AEDGTTLGVFSPKKLPHLVNLNEDPLMSECLIYYLKEGVTSVGRPEAEHRPDILLSGEAI  509
               : : ::  |      :  | :|: :|||| |||| :|:| |||:  |:| |||:|:||
Qy     480  IEKGF-VGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGGTRVGNVNQDTQAEIRLNGSKI  538

Db     510  LELHCEFINEDGNVTLTMKPNASCYINGKQVTTPTVLHTGSRVILGEHHVFRYNDPQEAR  569
               ||| |: : :|:  :: :||  |     :|:|:: ::| :|:|:|| ||:|:||:|||
Qy     539  LKEHCTFENVDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEAR  598

Db     570  QSRHN 574
               |::
Qy     599  AERQE 603
```

```
CC   -!- SUBUNIT: TRIMER OF A 115 KDA SUBUNIT AND TWO KINESIN-LIKE SUBUNITS
CC       OF 95 AND 85 KDA.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC       II SUBFAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L16993; AAA16098.1; -.
DR   HSSP; P17119; 3KAR.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
FT   DOMAIN        1    340    MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN      341    619    COILED COIL (BY SIMILARITY).
FT   DOMAIN      620    699    GLOBULAR (BY SIMILARITY).
FT   NP_BIND      97    104    ATP (POTENTIAL).
SQ   SEQUENCE  699 AA;   78697 MW;  7B3866111CB08190 CRC64;
```

```
Query Match       18.6%;   Score 1008;   DB 1;   Length 699;
Best Local Similarity   45.0%;   Pred. No. 3.40e-173;
Matches     182;   Conservative   86;   Mismatches 111;   Indels   25;   Gaps   15;

Db       6  SGNDNVRVVVCRPLNSKETGQGFKSVVKMDEMRG--TVQV-TNPNAP-SG-----EPPKS  57
               ||:|||||| :: :|:||   |:   :||: : || |  : : ||
Qy       2  SGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKA  61

Db      58  FTFD-T--VF---APG-AKQTDVNQTARPIVDAIIEGYNGTIFAYGQTGTGKTFTMEG-  109
               :|| :  | :  ||  |: :::|:||| ||  ||||||:||||||||::|:::|
Qy      62  FAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY  121

Db     110  VRSQPELRGIIPNSFAHIFGHIAKEQENVRPFLVRVSYLEIYNEEVKDLLGKDQQHRLEVK  169
               |:| ||::  |:|| |: |  :|:|| : : :::|:|||||:| | :|| | |:|
Qy     122  GKEHGVIPRICQDMFRRI-NELQKDK-NLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVR  179

Db     170  ERPDVGVYKDLSAFVVNNADDMDRIMTLGNKNRSVGATNMNESSRSHAIFTITL-ERS  228
               ||||| ||:|:| |  :: |: ::|||||:|:|:||:||||:||||:|:|:|||  ::
Qy     180  EHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKW  239

Db     229  DMGLDKEQHVRVGKLHMVDLAGSERQTKTGATGQRLKEATKINLSLTLGNVISSLVD--  286
               ::|| |::  | | |:||||||||| | |||| |||||:| :||||:|| :| :|
Qy     240  HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMS  299

Db     287  -GK--STH-IPYRNSKLTRLLQDSLGGNAKTVMCANIGPAEYNYDETISTLRYANRAKNI  342
               || ::|| :|||||| |||:| :: :||::::|| |||||::| ||: |||||||: |
Qy     300  SGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRI  359

Db     343  KNKAKINEDPKDALLREFQKEIEELKKQISESGEGLDDDEESGS  386
               ||:|||||| :: :| ||::|:|||::|
Qy     360  KNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGGGSGGGGSGGG  403
```

```
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF3A (MICROTUBULE PLUS END-DIRECTED KINESIN
DE   MOTOR 3A).
GN   KIF3A OR KIF3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93077686.
RA   Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M., Hirokawa N.;
RT   "Kinesin family in murine central nervous system.";
RL   J. Cell Biol. 119:1287-1296(1992).
CC   -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
CC       MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC       ACTIVITY IN VITRO.
CC   -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.
CC   -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
CC       TISSUE (MAINLY IN THE CEREBELLAR GRANULAR LAYER) WITHIN A SINGLE
CC       TYPE OF NEURONAL CELL.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC       II SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D12645; BAA02166.1; -.
DR   PIR; B44259; B44259.
DR   HSSP; P56536; 2KIN.
DR   MGD; MGI:107689; KIF3A.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
FT   DOMAIN        1    350       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN      351    586       COILED COIL (BY SIMILARITY).
FT   DOMAIN      587    701       GLOBULAR.
FT   NP_BIND     100    107       ATP (BY SIMILARITY).
FT   DOMAIN      442    445       POLY-GLU.
FT   DOMAIN      509    512       POLY-ARG.
SQ   SEQUENCE   701 AA;  80167 MW;  2405872DF2D85A29 CRC64;

Query Match             17.7%;  Score 962;  DB 1;  Length 701;
Best Local Similarity   41.4%;  Pred. No. 1.37e-163;
Matches  196;  Conservative  104;  Mismatches 137;  Indels 36;  Gaps 28;

Db         14 NVKVVVRCRPLNERE--K-SMCY-R-QAV-SV-DEMRGTI-TVHKTDSSN-EPPKTFTFD 64
Qy          6 NIKVVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAFAFD 65

Db         65 -T--VF---GPE-SKQLDVY-NLTARPIIDSVLEGYNGTIFAYGQTGTGKTFTMEGVRAV 116
Qy         66 RSYWSFDKNAPNYARQEDLFQDLGV-PLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-G- 122

Db        117 PGLRGVIPNSFAHIFGHIAKAEGDTRFL--VRVSYLEIYNEVRDLLGKDQTQRLEVKER 174
Qy        123 KE-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181

Db        175 PDVGVYIKDLSAYVNNADDMDRIMTLGHKNRSVGATNMNEHSSRSHAIFTITIECSEKG 234
Qy        182 PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFLITLTQKWHD 241

Db        235 VDGNMHV-RMGKLHLVDLAGSERQAKTGATGQRLKEATKINLSLSTLGNVISALVD---G 290
Qy        242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301
```

```
Db        291 K--STH-VPYRNSKLTRLLQDSLGGNSKTMMCANIGPADYNDETISTLRYANRAKNIKN 347
Qy        302 KQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFETLSTLRYADSAKRIKN 361

Db        348 KARINEDPKDALLRQFQKEIEELKKKLEE-GEEVSGSDISGSEEDDE-EGELGEDGEKRK 405
Qy        362 HAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGAGGSGGPVEESYPPDTPLEKQIVS 421

Db        406 KRRDQAGKKVSPDKMVE-M-QA-KIDEE-RKALETKLDMEEERNKARAELE 454
Qy        422 IQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEIHKEREAALE 474

RESULT  8
ID   KF3A_HUMAN     STANDARD;     PRT;    702 AA.
AC   Q9Y496;
DT   15-FEB-2000 (Rel. 39, Created)
DT   15-FEB-2000 (Rel. 39, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF3A (MICROTUBULE PLUS END-DIRECTED KINESIN
DE   MOTOR 3A).
GN   KIF3A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=RETINA;
RA   Wang S.Y., Bost-Usinger L., Hoang E., Frazer K.A., Reichelt S.,
RA   Burnside B.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
CC       MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC       ACTIVITY IN VITRO.
CC   -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC       II SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF041853; AAC72294.1; -.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
FT   DOMAIN        1    350       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN      351    587       COILED COIL (BY SIMILARITY).
FT   DOMAIN      588    702       GLOBULAR.
FT   NP_BIND     100    107       ATP (BY SIMILARITY).
FT   DOMAIN      443    446       POLY-GLU.
SQ   SEQUENCE   702 AA;  80337 MW;  24507D1EC6D540DB CRC64;

Query Match             17.7%;  Score 958;  DB 1;  Length 702;
Best Local Similarity   40.8%;  Pred. No. 9.35e-163;
Matches  193;  Conservative 107;  Mismatches 138;  Indels 35;  Gaps 27;

Db         14 NVKVVVRCRPLNERE--K-SMCY-K-QAV-SV-DEMRGTI-TVHKTDSSN-EPPKTFTFD 64
Qy          6 NIKVVVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAFAFD 65

Db         65 -T--VF---GPE-SKQLDVY-NLTARPIIDSVLEGYNGTIFAYGQTGTGKTFTMEGVRAI 116
Qy         66 RSYWSFDKNAPNYARQEDLFQDLGV-PLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-GK 123

Db        117 PELRGIIPNSFAHIFGHIAKAEGDTRFL-VRVSYLEIYNEEVRDLLGKDQTQRLAVKER 174
Qy        124 -E-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181
```

```
Db   175  PDVGVYIKDLSAYVVNNADDMDRIMTLGHKNRSVGATMNNEHSSRSHAIFTITIECSEKG 234
          |-|:  |:|:  | .| :|   ||||:::|||:::::::::|:
Qy   182  PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLQKWHD 241

Db   235  IDGNMHV-RMGKLHLVDLAGSERQAKTGATGQRLKEATKINLSLSTLGNVISALVD---G 290
          : .:  :||||||||||||   |:|  ||||  |:|:||| ||||
Qy   242  EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301

Db   291  K--STH-VPYRNSKLTRLLQDSLGGNSKTMMCANIGPADYNYDETISTLRYANRAKNIKN 347
          | :|||| |.  |||:::::::||||  |:||| |||| :| |||
Qy   302  KQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKN 361

Db   348  KARINEDPKDALLRQFQKEIEELKKKLEE-GEEISGSDISGSEEDDDEEGEVGEDGEKRK 406
          | :|||| |: |:  |::  | :|:  |:
Qy   362  HAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGAGGSGGPVEESYPPDTPLEKQIVS 421

Db   407  KRRIQIGKKKVSPDKMIE-M-QA-KIDEE-RKALETKLDMEEEERNKARAELE 455
          :: :: |: ::| . | :|:: | :|| |  |
Qy   422  IQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEIHKEREAALE 474
```

```
RESULT   9
ID   K121_STRPU    STANDARD;    PRT;    742 AA.
AC   P46871;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   KINESIN-II 95 KDA SUBUNIT (KRP-85/95 95 KDA SUBUNIT).
GN   KRP95.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;
OC   Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=EGG;
RX   MEDLINE; 94050179.
RA   Cole D.G., Chinn S.W., Wedaman K.P., Hall K., Vuong T.,
RA   Scholey J.M.;
RT   "Novel heterotrimeric kinesin-related protein purified from sea
RT   urchin eggs.";
RL   Nature 366:268-270(1993).
CC   -!- SUBUNIT: TRIMER OF A 115 KDA SUBUNIT AND TWO KINESIN-LIKE SUBUNITS
CC       OF 95 AND 85 KDA.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC       II SUBFAMILY.
```

```
DR   EMBL; U00996; AAA87393.1; -.
DR   HSSP; P56536; 2KIN.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
FT   DOMAIN       1    337  MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN     338    613  COILED COIL (BY SIMILARITY).
FT   DOMAIN     614    742  GLOBULAR (BY SIMILARITY).
FT   NP_BIND     95    102  ATP (POTENTIAL).
SQ   SEQUENCE   742 AA;  84202 MW;  47C40A367BAA77B5 CRC64;

Query Match             17.78%;  Score 962;  DB 1;  Length 742;
Best Local Similarity   44.58%;  Pred. No. 1.37e-163;
Matches   177;  Conservative  86;  Mismatches 105;  Indels 30;  Gaps 18;
```

```
Db     5  SAETVKVVVRCRPMNSKEISQGHKRIVEMDNKRG-LVEVTNP--KGP-PGEP--N-KSF 56
          ::|:|:|  ||| |::|  | :|   |.|| | :  :|:|
Qy     3  GGGNIKVVVRPPNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAF 62

Db    57  TFDTVYDWN-SKQIDLY---DETFRSL-V---ESVLQGFNGTIFAYGQTGKTFTMEGV 108
          :||  |:  :  ||  | :.| :|:|||||||
Qy    63  AFDRSY-WSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY 121

Db   109  RSNPELRGVIPNSFEHIFTHIARTQNQQFL---VRASYLEIYQEEIRDLLAKDQKKRLDL 165
          | ::|||| ||||  |. :|  | :|  : |||||||||||||
Qy   122  -GK-E-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKV 178

Db   166  KERPDTGVYVKDLSSFVTKSVKEIHVMTVGNNNRSVGSTMMNEHSSRSHAIFIITIECS 225
          :|: ||:|:  |::  |::   |: |: ||||||||:||||||||
Qy   179  REHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQK 238

Db   226  ELGVDGENHI-RVGKLNLVDLAGSERQAKTGATGDRLKEATKINLSLSALGNVISALVD- 283
          :|::|||||||   |:| ||| |:|:|  ||||
Qy   239  WHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADM 298

Db   284  --GK--SSH-IPYRDSKITRLLQDSLGGNAKTVMVANMGPASYNFDETITTLRYANRAKN 338
          : |||||||||  | :|  | ::
Qy   299  SSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKR 358

Db   339  IKNKPKINEDPKDALLREFQEISRLKQALDKKGPSDG 376
          |||  | :|| ::| |:  :|:| :  |  |
Qy   359  IKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGG 396
```

```
RESULT  10
ID   KINL_LEICH    STANDARD;    PRT;    955 AA.
AC   P46865;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   KINESIN-LIKE PROTEIN K39 (FRAGMENT).
GN   KIN.
OS   Leishmania chagasi.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MHOM/BR/82 / ISOLATE BA-2;
RX   MEDLINE; 93133867.
RA   Burns J.M. Jr., Shreffler W.G., Benson D.R., Ghalib H.W., Badaro R.,
RA   Reed S.G.;
RT   "Molecular characterization of a kinesin-related antigen of
RT   Leishmania chagasi that detects specific antibody in African and
RT   American visceral leishmaniasis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).
CC   -!- DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
```

```
DR   EMBL; L07879; AAA29254.1; -.
DR   HSSP; P17119; 3KAR.
DR   PFAM; PF00225; kinesin; 1.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Repeat.
FT   DOMAIN       1    399  MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN     426   >955  COILED COIL (POTENTIAL).
FT   NP_BIND    122    129  ATP (POTENTIAL).
FT   DOMAIN     704   >955  7 X 39 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT     704    742  1.
FT   REPEAT     743    781  2.
```

```
FT   REPEAT         782     820       3.
FT   REPEAT         821     859       4.
FT   REPEAT         860     898       5.
FT   REPEAT         899     937       6.
FT   NON_TER        938    >955       7 (PARTIAL).
SQ   SEQUENCE    955 AA;  106168 MW;  8CA76815BE84C6E9 CRC64;

Query Match             17.6%;  Score 953;  DB 1;  Length 955;
Best Local Similarity   43.5%;  Pred. No. 1.03e-161;
Matches  177;  Conservative  90;  Mismatches 111;  Indels  29;  Gaps 18;

Db    13 VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVTVKVLGGSNNSGAAESMGTARRVAQD   72
         :|| |||||||  |:  |:|   |   |   :| |:     :  :  :
Qy     7 IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA   61

Db    73 FQFDHVFWSVETPDACGAIPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM  132
         | |:||  :| :||   |   :||||||||||||:|:|
Qy    62 FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGSYSM  118

Db   133 MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIIVELGYVEVYNERVSDLLGKRK  192
          || :  | :| ||||  |  |  | |:||||  |||| |
Qy   119 MG------Y-GKEHGVIPRICQDMFRRINELQKDKNLTCYVEVSLEIYNERVRDLL-NP-  170

Db   193 KGVKGGGEEVVVDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR  251
          :  ||  : ||||||| |:  :|| :|  |  |:::  : :|:||:|:
Qy   171 -STKG---NL--KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET  223

Db   252 SSRSHAIIMLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSQVEGQQFKEAT  311
          |||||::    ||| | :   | :::||  ::||||||||||| |::    :::
Qy   224 SSRSHAVFTLTQKW----HDEETKMDTEVAKISLVDLAGSERATSTGATGARLKEGA  279

Db   312 HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP  371
          ||| |:|||| || ||  :||| :|:|||||||||:||||||||||||||:||:|
Qy   280 EINRSLSTLGRVIAALADMSS-GKOKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP  338

Db   372 SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR  418
          : :|||||||||  :  |:  |   |||| || |:|
Qy   339 ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR  385

RESULT 11
ID   KF3B_HUMAN     STANDARD;      PRT;   747 AA.
AC   O15066;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF3B (MICROTUBULE PLUS END-DIRECTED KINESIN
DE   MOTOR 3B) (KIAA0359) (HH0048).
GN   KIF3B.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 97349984.
RA   Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VII.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 4:141-150(1997).
CC   -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
CC       MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC       ACTIVITY IN VITRO (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC       II SUBFAMILY.
```

```
DR   EMBL; AB002357; BAA20815.1; -.
DR   HSSP; P56536; 2KIN.
DR   MIM; 603754; -.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
FT   DOMAIN         1     345       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN       346     579       COILED COIL (BY SIMILARITY).
FT   DOMAIN       580     747       GLOBULAR.
FT   NP_BIND       96     103       ATP (POTENTIAL).
FT   DOMAIN       386     393       POLY-GLY.
FT   DOMAIN       394     406       POLY-GLU.
FT   DOMAIN       723     730       POLY-SER.
SQ   SEQUENCE    747 AA;  85125 MW;  97FA4573AFA87023 CRC64;

Query Match             17.4%;  Score 945;  DB 1;  Length 747;
Best Local Similarity   47.0%;  Pred. No. 4.78e-160;
Matches  164;  Conservative  72;  Mismatches  99;  Indels  14;  Gaps 9;

Db    37 KLGQVSVKNPKGTAHEMPKTFFDAVYDWNAKQFELYDETFRPLVDSVLQGFNGTIFAYG   96
         |||     :||   ||   :||  ||| :|  :||:    |:||||| ||||||
Qy    50 KSGKTIMDGPKAFAFDRS-YWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYG  108

Db    97 QTGTGKTYTMEGIRGDPEKRGVIPNSFDHIFTHISRSQNQYL---VRASLEIYQEIR   153
         |||:||:|:||  |   | ||||    |::|  |:  | |||   |:||||||::| |
Qy   109 QTGSGKSYSMGY-GK-E-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVR   165

Db   154 DLLSKDQTKRLELKERPDTGVYVKDLSSFVTKSVKEIEHVMNVGNQNRSVGATNMNEHS   213
         ||| :     |  :|:||:|:| |   :| |:|:|||| :|:| | |::  ||||||: |
Qy   166 DLLNPSTKGNLKVREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSS  225

Db   214 RSHAIFVITI-ECSEVGLDGENHIRVGKLNLVDLAGSERQAKTGAQGERLKEATKINLSL  272
         |||:| :    | :  ::|  :  ||| ||:|||||||||:||:|:  |||  :||||
Qy   226 RSHAVFTLTLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSL  285

Db   273 SALGNVISALVD--GK--STH-IPYRDSKLTRLLQDSLGGNAKTVMVANVGPASYNVEE  326
         : || ||:|  | ||   |:  ||||| ||| ||:|||||:   |:|:  | |  |  |
Qy   286 STLGRVIAALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEE  345

Db   327 TLTTLRYANRAKNIKNKPRVNEDPKDALLREFQEIARLKAQLEKRSIG  375
         ||:||||| :|:|||| |||| ||: |:|| :|||:||:: ||  :|
Qy   346 TLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKIQSGSGGG  394

RESULT 12
ID   KF3B_MOUSE     STANDARD;      PRT;   747 AA.
AC   Q61771;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF3B (MICROTUBULE PLUS END-DIRECTED KINESIN
DE   MOTOR 3B).
GN   KIF3B.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=BRAIN;
RX   MEDLINE; 96032268.
RA   Yamazaki H., Nakata T., Okada Y., Hirokawa N.;
RT   "KIF3A/B: a heterodimeric kinesin superfamily protein that works as a
RT   microtubule plus end-directed motor for membrane organelle
RT   transport.";
RL   J. Cell Biol. 130:1387-1399(1995).
CC   -!- FUNCTION: MICROTUBULE-BASED ANTEROGRADE TRANSLOCATOR FOR
CC       MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
```

```
CC      ACTIVITY IN VITRO.
CC  -!- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.
CC  -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC      II SUBFAMILY.
CC  --------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------
DR  EMBL; D26077; BAA05070.1; -.
DR  HSSP; P55536; 2KIN.
DR  MGD; MGI:107688; KIF3B.
DR  PFAM; PF00225; kinesin; 1.
DR  PRINTS; PR00380; KINESINHEAVY.
DR  PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR  PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW  Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
FT  DOMAIN        1    345       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT  DOMAIN      346    579       COILED COIL (BY SIMILARITY).
FT  DOMAIN      580    747       GLOBULAR.
FT  NP_BIND      96    103       ATP (POTENTIAL).
FT  DOMAIN      386    393       POLY-GLY.
FT  DOMAIN      394    405       POLY-GLU.
FT  DOMAIN      723    730       POLY-SER.
SQ  SEQUENCE   747 AA;  85288 MW;  FA369A4190EC8B47 CRC64;

Query Match             17.4%;  Score 942;  DB 1;  Length 747;
Best Local Similarity   46.7%;  Pred. No. 2.02e-159;
Matches  163;  Conservative  73;  Mismatches  99;  Indels  14;  Gaps  9;

Db   37  KLGQVSVKNPKGTSHEMPKTFTFDAVYDWNAKQFELYDETFRPLVDSVLQGFNGTIFAYG  96
Qy   50  KSGKTIMDGPKAFAFDRS-YWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYG 108

Db   97  QTGTGKTYTMEGVRGDPEKRGVIPNSFDHIFTHISRSONQOYL---VRASYLEIYQEEIR 153
Qy  109  QTGSGKSYSMMGY-GK-E-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVR 165

Db  154  DLLSKDQTKRLELKERPDTGVYVKDLSSFVTKSVKEIEHVMNVGNQNRSVGATNMNEHSS 213
Qy  166  DLLNPSTKGNLKVREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSS 225

Db  214  RSHAIFVITI-ECSEVGLDGENHIRVGKLNLVDLAGSERQAKTGAQGERLKEATKINLSL 272
Qy  226  RSHAVFTLTLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSL 285

Db  273  SALGNVISALVD---GK--STH-IPYRDSKLTRLLQDSLGGNAKTVMVANVGPASYNVEE 326
Qy  286  STLGRVIAALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEE 345

Db  327  TLTLRTANRAKNIKNKPRVNEDPKDALLREFQEEIARLKAQLEKRSIG 375
Qy  346  TLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGG 394

RESULT  13
ID  KL68_DROME   STANDARD;   PRT;   784 AA.
AC  P46867;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  KINESIN-LIKE PROTEIN KLP68D.
GN  KLP68D OR KLP5.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 95050960.
RA  Pesavento P.A., Stewart R.J., Goldstein L.S.B.;
RT  "Characterization of the KLP68D kinesin-like protein in Drosophila:
RT  possible roles in axonal transport.";
RL  J. Cell Biol. 127:1041-1048(1994).
RN  [2]
RP  SEQUENCE OF 220-342 FROM N.A.
RX  MEDLINE; 92020874.
RA  Stewart R.J., Pesavento P.A., Woerpel D.N., Goldstein L.S.B.;
RT  "Identification and partial characterization of six members of the
RT  kinesin superfamily in Drosophila.";
RL  Proc. Natl. Acad. Sci. U.S.A. 88:8470-8474(1991).
CC  -!- FUNCTION: PLUS-END DIRECTED MICROTUBULE MOTOR THAT MAY BE USED FOR
CC      ANTEROGRADE AXONAL TRANSPORT AND COULD CONCEIVABLY MOVE CARGOES IN
CC      FLY NEURONS DIFFERENT THAN THOSE MOVED BY KINESIN HEAVY CHAIN OR
CC      OTHER PLUS-END DIRECTED MOTORS.
CC  -!- TISSUE SPECIFICITY: EXPRESSED PRIMARILY IN THE CENTRAL NERVOUS
CC      SYSTEM AND IN A SUBSET OF THE PERIPHERAL NERVOUS SYSTEM DURING
CC      EMBRYOGENESIS.
CC  -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC      II SUBFAMILY.
CC  --------------------------------------------------------------------
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  --------------------------------------------------------------------
DR  EMBL; U15974; AAA69929.1; -.
DR  EMBL; M74431; AAA28658.1; -.
DR  HSSP; P55536; 2KIN.
DR  FLYBASE; FBgn0004381; Klp68D.
DR  PFAM; PF00225; kinesin; 1
DR  PRINTS; PR00380; KINESINHEAVY.
DR  PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR  PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW  Motor protein; Microtubules; ATP-binding; Coiled coil.
FT  DOMAIN       16    275       MECHANOCHEMICAL (MOTOR).
FT  DOMAIN      351    385       COILED COIL (POTENTIAL).
FT  DOMAIN      426    582       COILED COIL (POTENTIAL).
FT  NP_BIND     106    113       ATP (POTENTIAL).
FT  CONFLICT    220    221       SS -> TC (IN REF. 2).
FT  CONFLICT    338    342       GSRAK -> VRGQV (IN REF. 2).
SQ  SEQUENCE   784 AA;  88193 MW;  94BB9BADF072DFC0 CRC64;

Query Match             16.5%;  Score 897;  DB 1;  Length 784;
Best Local Similarity   46.5%;  Pred. No. 4.60e-150;
Matches  152;  Conservative  83;  Mismatches  72;  Indels  20;  Gaps  14;

Db   63  QRKVFTYDAAYDASATQTTLYHEVVFPLVSSVLEGFNGCIFAYGQTGTGKTFTMEGVRGN 122
Qy   65  DRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGVNNCIFAYGQTGSGKSYSMMGY-GK 123

Db  123  DELMGIIPRTFEQIWLHIN--RTE-NFQFLVDVSYLEIYMEELRDLLKPNSK-HLEVRER 178
Qy  124  EH--GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181

Db  179  -GSGVYVPNLHAINCKSVEDMIKVMQVGNKNRTVGFTMNEHSSRSHAIFMIKI-EM-CD 235
Qy  182  PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241

Db  236  TET-NTIKVGKLNLIDLAGSERQSKTGASAERLKEASKINLALSSLGNVISALAE-SS- 291
Qy  242  EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301

Db  292  -P--H-VPYRDSKLTRLLQDSLGGNSKTIMIANIGPSNVYNETLLTTLRYGSRAKSIQN 346
Qy  302  KQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKN 361

Db  347  QPIKNEDPQDAKLKEYQEEIERLK-RL 372
```

```
Qy    362 HAVVNEDPNARMIRELKEELAQLRSKL 388

RESULT 14
ID   FL10_CHLRE     STANDARD;      PRT;    786 AA.
AC   P46869;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   KINESIN-LIKE PROTEIN FLA10 (KHP1 PROTEIN).
GN   FLA10.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=137;
RX   MEDLINE; 94299638.
RA   Walther Z., Vashishtha M., Hall J.L.;
RT   "The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous
RT   protein.";
RL   J. Cell Biol. 126:175-188(1994).
CC   -!- FUNCTION: PROBABLY INVOLVED IN FLAGELLAR ASSEMBLY AND MAINTENANCE.
CC       MAY PLAY A ROLE IN FLAGELLAR SYNTHESIS.
CC   -!- TISSUE SPECIFICITY: FLAGELLAR AXONEME.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC       II SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L33697; AAA21738.1; -.
DR   HSSP; P17119; 3KAR.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
FT   DOMAIN       1    358       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN     367    687       COILED COIL (POTENTIAL).
FT   DOMAIN     688    786       GLOBULAR (POTENTIAL).
FT   NP_BIND     97    104       ATP (POTENTIAL).
FT   DOMAIN     388    391       POLY-GLY.
FT   DOMAIN     705    714       POLY-GLY.
FT   DOMAIN     756    759       POLY-ASP.
SQ   SEQUENCE   786 AA;  86671 MW;  F90962203EB79F1B CRC64;

Query Match             16.0%;  Score 870;  DB 1;  Length 786;
Best Local Similarity   40.0%;  Pred. No. 1.84e-144;
Matches  187;  Conservative  105;  Mismatches 138;  Indels  38;  Gaps  22;

Db     6 GGSESVKVVVRCRPLNGEKADGRSRIVDMD----V-DA--G-QVKVRNPKADASEPPKA 57
          : :: :||||||  || ||:::|                | : :|:: |||  ::||
Qy     2 SGGGNIKVVVVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKA 61

Db    58 FTFDQVY---DWNCQQ--R--DVFDITARPLIDSCIEGYNGTIFAYGQTGKSHTMEG- 109
          | :|| | |          | : :|:: | ||||||||||:||| ||:|
Qy    62 FAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY 121

Db   110 KDEPPELRGLIPNTFRYVFEIIARDSGTKEFLVRSSYLEIYNEEVRDLLGKDHSKKMELK 169
          | || |  | |  ||: ||::| |:|  | : || ||||||  |:||  |:::  ::
Qy   122 GKEHGVIPRICQDMFRRINEL-QKDKNLT-CTVEVSYLEIYNERVRDLLNPSTKGNLKVR 179

Db   170 ESPDRGVYVKDLSQFVCKNYEEMNKVLLAGKDNRQVGATLMNQDSSRSHSIFTITIECIE 229
          |  |||:    ||:| |      |  | :||  |: |: |  |  | :||||::||: :
Qy   180 EHPSTGPYVEDLAKILVVRSFQEIENLMDEGNKARPVAATNMNETSSRSHAVFTLTL--TQ 237

Db   230 KLESAAAQKPGAKKDDSNHVRVGKLNLVDLAGSERQDKTGATGDRLKEGIKINLSLTALG 289
```

```
Qy   238 KWHDE-E---TKMD-TE--KVAKISLVDLAGSERATSGATGARLKEGAEINRSLSTLG 289
         :|:|| :: :        :|::|::||||||| ||:| |: |||:||||||| :::||

Db   290 NVISALVD--GK--SGH-IPYRDSKLTRLLQDSLGGNTKTVMVANIGPADWNYDETMST 343
         ||:|| :|   ||        :|||||:|||:|||||:|:|| :| :|||:||:::||

Qy   290 RVIAALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLST 349
         ||||| |   ||: || |  :||||:|| |||||||| | :|:|||||||:|||::||

Db   344 LRYANRAKNIQNKPKINEDPKDAMLRQFQEEIKKLKEQLAARAAGGGPI-TM-PSGGGS 401
         |||: || ||||:|| | |  :||:|||||||||||| || :|: ||:  || |||

Qy   350 LRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESY 409

Db   402 PTQKIVER-TEEVDPDIDAIKAQMRAELEAKMK-SDISTEALDKAREE 447
         | : :     || ||| || :||:|| | :      : :: || :||

Qy   410 PPDTPLEKQIVSIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEE 457


RESULT 15
ID   KIF4_MOUSE     STANDARD;      PRT;   1231 AA.
AC   P33174;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF4.
GN   KIF4 OR KNS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=BRAIN;
RX   MEDLINE; 95014709.
RA   Sekine Y., Okada Y., Noda Y., Kondo S., Aizawa H., Takemura R.,
RA   Hirokawa N.;
RT   "A novel microtubule-based motor protein (KIF4) for organelle
RT   transports, whose expression is regulated developmentally.";
RL   J. Cell Biol. 127:187-201(1994).
RN   [2]
RP   SEQUENCE OF 91-240 FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93077686.
RA   Aizawa H., Sekine Y., Takemura R., Zhang Z., Nangaku M.,
RA   Hirokawa N.;
RT   "Kinesin family in murine central nervous system.";
RL   J. Cell Biol. 119:1287-1296(1992).
CC   -!- FUNCTION: THE KINESIN FAMILY MAY RECOGNIZE, TRANSPORT AND POSITION
CC       THEIR SPECIFIC CARGOES IN A SINGLE TYPE OF NEURONAL CELL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT SPLEEN AND JUVENILE BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D12646; BAA02167.1; -.
DR   PIR; D44259; D44259.
DR   HSSP; P17119; 3KAR.
DR   MGD; MGI:108389; KIF4.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
FT   DOMAIN       1    350       MECHANOCHEMICAL (MOTOR).
FT   DOMAIN     351   1000       COILED COIL (BY SIMILARITY).
FT   DOMAIN    1001   1231       GLOBULAR.
FT   NP_BIND     88     95       ATP (POTENTIAL).
FT   CONFLICT   112    112       I -> S (IN REF. 2).
SQ   SEQUENCE  1231 AA;  139551 MW;  F34F2C2D1158FE4 CRC64;
```

```
Query Match            15.6%;   Score 846;  DB 1; Length 1231;
Best Local Similarity 47.8%;   Pred. No. 1.71e-139;
Matches  163;  Conservative  70;  Mismatches  82;  Indels  26;  Gaps  18;

Db   46 DKSF-TYDFVFDPS-TEQEEVFNTAVAPLIKGVFKGYNATVLAYGQTGSGKTYSMGGAYT 103
           |:|: ::|  |: : ||  :||    ||||| :||||||||||||:|||| |   |
Qy   65 DRSVWSFD-KNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG-Y- 121

Db  104 AEQEHDSAIGVIPRVIQLLFKEIN--KKS-DFEFTLKVSYLEIYNEEILDLLCSSREKAT 160
           : || |||||:|:||::|:|| |  :||  : |:||||||| : ||| ::  :|:
Qy  122 G-KEH----GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLL-NPSTKGN 175

Db  161 QINIREDPKEGIKIVGLTEKTVLVASDTVSCL-EQGNNSRTVASTAMNSQSSRSHAIFTI 219
           ::||  |   | |   |: |   : : : |  : ::|::||||:| || ||||||::||:
Qy  176 -LKVREHPSTGPYVEDLA-KLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTL 233

Db  220 SIEQR-KKND-KNSSFR-SKLHLVDLAGSERQKKTKAEGDRLREGININRGLLCLGNVIS 276
           ::|:  : :|   :: : ||||||||||||    |  || |:|| :|||:|  ||  |
Qy  234 TLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIA 293

Db  277 ALGD-----DKKGNFVPYRDSKLTRLLQDSLGGNSHTLMIACVSPADSNLEETLNTLRYA 331
           ||:|     :|| |||||||| ||:|| |||||| |||  :|||| |:|||| ::||||
Qy  294 ALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYA 353

Db  332 DRARKIKNKPIINIDPQAAELNHLKQQVQQLQILLLQAHGG 372
           |:|:||| | ||| :  :::: || ||: | |: ||
Qy  354 DSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKL-QSSGG 393


Search completed: Mon Aug 21 15:40:50 2000
Job time : 30 secs.
```

THIS PAGE BLANK (USPTO)

```
 __  __ ____  ____  ____   ___ _   _
|  \/  |  _ \/ ___||  _ \ / __| | | |
| |\/| | |_) \___ \| |_) | |  | |_| |
| |  | |  __/ ___) |  _ <| |__|  _  |
|_|  |_|_|   |____/|_| \_\\___|_| |_|  (TM)
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 21 15:41:08 2000; MasPar time 60.85 Seconds
           893.265 Million cell updates/sec

Tabular output not generated.

Title: >US-09-235-416-1
Description: (1-784) from US09235416A.pep
Perfect Score: 5422
Sequence: 1 MSGGGNIKVVVRVRPFNARE........ELRQQQAQMEEALKTAKQEF 784

Scoring table: PAM 150
                  Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database: sptrembl2
                  1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
                  5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
                  9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
                  13:sp_vertebrate 14:sp_virus

Statistics: Mean 52.569; Variance 102.972; scale 0.511

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2215	40.9	1770	11	Q92119	KIF1B-BETA.	0.00e+00
2	2051	37.8	1103	4	075186	KIAA0706 PROTEIN.	0.00e+00
3	2006	37.0	1097	11	035787	KINESIN-RELATED PROTEI	0.00e+00
4	1971	36.4	1584	5	Q18778	C. ELEGANS UNC-104 KIN	0.00e+00
5	1858	34.3	1816	11	Q9WVE5	KIF1B MAJOR ISOFORM.	0.00e+00
6	1849	34.1	689	11	08B658	KINESIN-LIKE PROTEIN K	0.00e+00
7	1597	29.5	1921	5	001349	KINESIN-73.	1.37e-300
8	1308	24.1	1648	4	Q15058	KIAA0042 PROTEIN.	9.45e-239
9	1291	23.8	928	5	Q09997	PUTATIVE KINESIN-LIKE	3.96e-235
10	1124	20.7	1576	5	Q20888	SIMILAR TO KINESIN-REL	1.18e-199
11	958	17.7	702	4	Q9Y496	KINESIN FAMILY MEMBER	1.22e-164
12	904	16.7	744	13	093478	KINESIN LIKE PROTEIN 3	2.57e-153
13	857	15.8	1121	5	018389	KINESIN-LIKE PROTEIN,	1.71e-143
14	857	15.8	1121	5	002030	KINESIN-LIKE PROTEIN.	1.71e-143
15	854	15.8	1121	5	018866	KINESIN LIKE PROTEIN A	7.23e-143
16	850	15.7	1121	5	018390	CHROMATIN ASSOCIATED K	4.94e-142
17	822	15.2	402	4	Q9Y6V4	KIF3 (FRAGMENT).	3.36e-136
18	823	15.2	928	5	Q90640	CHROMOKINESIN.	2.08e-136
19	825	15.2	1225	13	091784	KINESIN-LIKE PROTEIN 1	7.98e-137
20	796	14.7	929	3	P7B718	KINESIN.	8.54e-131
21	791	14.6	2954	13	042263	KINESIN-RELATED PROTEI	9.32e-130
22	777	14.3	968	3	P87199	KINESIN MOTOR PROTEIN.	7.47e-127
23	777	14.3	1212	13	Q9XX29	HYPOTHETICAL 135.8 KD	7.47e-127
24	736	13.6	1388	13	Q91785	KLP2 PROTEIN.	2.28e-118
25	734	13.5	150	4	095068	MICROTUBULE-BASED MOTO	5.89e-118
26	732	13.5	1254	5	094463	K7 KINESIN-LIKE PROTEI	1.52e-117
27	720	13.3	814	5	P91945	KINESIN LIKE PROTEIN 6	4.56e-115
28	704	13.0	797	11	08B657	KINESIN-LIKE PROTEIN K	9.01e-112
29	703	13.0	1027	11	Q922F9	KINESIN HEAVY CHAIN.	1.45e-111
30	697	12.9	151	13	Q91112	KINESIN HEAVY CHAIN FK	2.48e-111
31	701	12.9	784	4	059751	PUTATIVE KINESIN-LIKE	3.73e-111
32	696	12.8	160	11	035063	MOTOR DOMAIN OF KIF13B	3.99e-110
33	695	12.8	160	11	035062	MOTOR DOMAIN OF KIF13A	6.40e-110
34	696	12.8	956	11	Q922F8	KINESIN HEAVY CHAIN.	3.99e-110
35	696	12.8	957	4	060282	KIAA0531 PROTEIN.	3.99e-110
36	692	12.8	1083	5	045935	Y43F4B.6 PROTEIN.	2.65e-109
37	688	12.7	151	13	Q91116	KINESIN HEAVY CHAIN FK	1.76e-108
38	689	12.7	1130	5	Q19633	F20C5.2 PROTEIN.	1.10e-108
39	682	12.6	987	10	081635	KINESIN-LIKE HEAVY CHA	3.00e-107
40	675	12.4	1014	11	Q92110	KINESIN-RELATED MITOTI	8.20e-106
41	655	12.1	773	13	P79805	MICROTUBULE-BASED MOTO	1.03e-101
42	656	12.1	1200	7	097277	PFC0860W PROTEIN.	6.41e-102
43	648	12.0	687	4	075299	MICROTUBULE-BASED MOTO	2.78e-100
44	651	12.0	1893	5	017887	T01G1.1 PROTEIN.	6.76e-101
45	626	11.5	932	5	Q17753	SIMILAR TO KINESIN-LIK	8.61e-96

ALIGNMENTS

```
RESULT    1
ID   Q92119              PRELIMINARY;     PRT;   1770 AA.
AC   Q92119;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KIF1B-BETA.
GN   KIF1B.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR MOUSE;
RA   NAKAGAWA T., HIROKAWA N.;
RT   "Identification and characterization of a new kinesin superfamily
RT   KIF1b-beta.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB023656; BAA75243.1; -.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE 1770 AA;  198851 MW;  4F752FFC CRC32;

Query Match             40.9%;  Score 2215;  DB 11;  Length 1770;
Best Local Similarity   55.3%;  Pred. No. 0.00e+00;
Matches  340;  Conservative  131;  Mismatches 115;  Indels  29;  Gaps  20;

Db    1 MSGAS-VKVAVRVRPFPNSRETSKESKCIIQMQGNSTSIINP-----K--NP-K---EAPK 48
          ||| :  :| ||||||||         ::|:|                K  NP K   ::|
Qy    1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNGTILTPPPGAEEKARKSGKTIMDGPK 60

Db   49 SFSPDYSYWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNVCIFAYGQTGAGKSYTMM 108
          :| |||||| :|| |     ||||| : :   :|| ||  :::|||||||||| |||| ||
Qy   61 AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db  109 GKQESQAGIIPQLCEELFEKIND-NCNEEMSYSVEVSYMEIYCERVRDLLNPKNKGNLR 167
          | :     GVIP: |  :   :: ||:    |:VEVSY EIY ERVRDLLNP KGNL:
Qy  120 GYGKE-H-GVIPRICQDMPRRINELQKDKNLTCVEVSYLEIYNERVRDLLNPSTKGNLK 177

Db  168 VREHPLLGPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATMNETSSRSHAVFTIVFTQ 227
          VREHP  GPYVEDL:KL V S  :I :LMD GNKARTVAATMNETSSRSHAVFT: :TQ
Qy  178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATMNETSSRSHAVFTLTLTQ 237

Db  228 KKQDPETNLSTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE 287
```

```
Qy  238  KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD  297
         :| ||::: ||||:|||||||||||||:::||||| ||||||||||:: |::||||||:
Db  288  VSK-KKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRYADRAK   346
Qy  298  MSSGKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK   357
Db  347  QIKCNAVINEDPNAKLVRELKEEVTRLKDLLRAQGLGDI-IDTSMGSLTSS-PSSCSLNS  404
Qy  358  RIKNHAVVNEDPNARMIRELKELEALAQLRSKL-QS-SGGGGAGGSGSGGAGGSGPVEES
Db  405  QVGLTSVTSIQERIMSTPGGEEAIERLKESEKIIAELNETWEEKLRKTEAIRMEREALLA  464
Qy  418  QI-V-SIQQPDATVKKMSKAE-IVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALE  474
Db  465  EMGVAIREDGGTLGVFSPKKTPHLVNLNEDPLMSECLLYYIKDGITRVGQADAERRQDIV  524
Qy  475  ELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIR  532
Db  525  LSGAHIKEEHCLFRSERSNTGEVIVTLEPCERSETYVNGKRVAHPVQLRSGNRIIMGKNH  584
Qy  533  LNGSKILKEHCTF-E--NVDNV-VTIVPNEKAAVMVNGVRIDKPTPRLRSGVRIIILGDFH  587
Db  585  VFRFNHPEQARAERE  599
Qy  588  IFRFNHPEEARAERQ  602
```

```
RESULT    2
ID   O75186        PRELIMINARY;    PRT;   1103 AA.
AC   O75186;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KIAA0706 PROTEIN.
GN   KIAA0706.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE; 98403880.
RA   ISHIKAWA K., NAGASE T., SUYAMA M., MIYAJIMA N., TANAKA A., KOTANI H.,
RA   NOMURA N., OHARA O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
DR   EMBL; AB014606; BAA31681.1; -.
DR   HSSP; P17119; 3KAR.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE  1103 AA;  122946 MW;  F0FBC5F0 CRC32;

Query Match    37.8%;
Best Local Similarity  54.3%;   Pred. No. 0.00e+00;
Matches  340;  Conservative  127;  Mismatches  127;  Indels  42;  Gaps  28;

Db     1  MAGAS-VKVAVRVRPFNARETSQDAKCVVSMQGNTTSIINP----K-Q-S-K--DAPK   48
          ||||| :|:||||||||||:|  |||:| |:||: :|:|        | | |   | |
Qy     1  MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK   60

Db    49  SFTFDYSYWSHTSTEDPQFASQQQVYRDIGEEMLLHAFEGYNVCIFAYGQTGAGKSYTMM  108
          :| |||||||: :|| |:|| ||||::|| ||:: :|| |||:|||||||||:|||||||
Qy    61  AFAFDRSYWSF-DKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db   109  GRQEPGQOGIVPQLCEDLFSRVSENQSAQ-LSYSVEVSMEIYCERVRDLLNPKSRGSLR  167
          |:|: | || ||::||::|||: :|| || |: :||||:||:||||||||||:::||:|
Qy   120  G-YGK-EHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK  177
```

```
Db  168  VREHPILGPYVQDLSKLAVTSYADIADLMDCGNKARTVAATNMNETSSRSHAVFTIVFTQ  227
         |||||  |||:|| |||  :|::|| :|||| ||||||||||||||||||||||| :| |
Qy  178  VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ  237
Db  228  RCHDQLTGLDSEKVKISLVDLAGSERADSSGARGMRLKEGANINKSLTTLGKVISALAD   287
Qy  238  KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD  297
Db  288  MQS-KKRKSDFIPYRDSVLTWLKENLGGNSRTAMIAALSPADINYEETLSTLRYADRTK  346
Qy  298  MSSGKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK   357
Db  347  QIRCNAIINEDPNARLIRELQEEVARLRELLMAQGLSASALEGLKTEEGSVRGALPAVSS  406
Qy  358  RIKNHAVVNEDPNARMIRELKEELEALAQLRSKL--QS-SGGG-GG--GAGGSG-GPVEE-SY  409
Db  407  PPAPVSPSSPPTHNGELEPSFSPNTESQIGPEEAMERLQETEKIIAELNETWEEKLRKTE  466
Qy  410  PPD--TPLE--KQIVSIQQP-DA-TVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTE  463
Db  467  ALRMEREALLAEMGVAVREDGGTVGVFSPKKTPHLVNLMEDPLMSECLLYHIKDGVTRVG  526
Qy  464  EIHKEREAALEELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVG  521
Db  527  QV--DM--DIKLTGQFIREQHCLFRSIPQPDGEVVTLEPCEGAETYVNGKLVTEPLVLK  582
Qy  522  NVNQDTQAEIRLNGSKILKEHCTFENV---D-NVV-TIVPNEKAAVMVNGVRIDKPTRLR  576
Db  583  SGNRIVMGKNHVVFRFNHPEQARLERE  608
Qy  577  SGYRIILGDFHIFRFNHPEEARAERQ  602
```

```
RESULT    3
ID   O35787        PRELIMINARY;    PRT;   1097 AA.
AC   O35787;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KINESIN-RELATED PROTEIN.
GN   KIF1D.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR;
RA   ROGERS K.R., GRIFFIN M., BROPHY P.J.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ000696; CAA04248.1; -.
DR   HSSP; P17119; 3KAR.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
SQ   SEQUENCE  1097 AA;  122333 MW;  90F0DC3A CRC32;

Query Match    37.0%;
Best Local Similarity  53.8%;   Pred. No. 0.00e+00;   DB 11;   Length 1097;
Matches  336;  Conservative  126;  Mismatches  121;  Indels  41;  Gaps  27;

Db     1  MAGAS-VKVAVRVRPFNARETSQDAKCVVSMQGNTTSIINP----K-Q-S-RMFL--K   48
          ||||| :|:||||||||||:|  |||:|:|:||: :|:|        | | | ||||  |
Qy     1  MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK   60

Db    49  A-SFDYSYWSHTSVEDPQFASQQQVYRDIGEEMLLHAFEGYNVCIFAYGQTGAGKSYTMM  107
          | :|||||:|||: :||| ||| ||||||||:||||||:||||||||||||:|||||||
Qy    61  AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db   108  GRQEPGQOGIVPQLCEDLFSRVNVNQSAQ-LSYSVEVSMEIYCERVRDLLNPKSRGSLR  166
          |:|::| || ||:||||:|||: ::|:|| |||||||||:|||||||||||||| :|:|
Qy   120  G-YGK-EHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK  177
```

```
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE   1584 AA;  179649 MW;  75E81D7D CRC32;

Query Match             36.4%;  Score 1971;  DB 5;  Length 1584;
Best Local Similarity   52.6%;  Pred. No. 0.00e+00;
Matches 318; Conservative 122; Mismatches 127; Indels 38; Gaps 20;

Db      2 SSVKVAVRVRPFNQREISNTSKCVLQVNGNTTIN---G-HSINKE-N-F-----SFNF  49
          ::||||||||| ||||| |:|  |:|:|::||||    |  | :  : :|      :||
Qy      5 GNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAFAF  64

Db     50 DHSYWSFARNDPHFITQKQYYEELGVEMLEHAFEGYNVCIFAYGQTGSGKSYTMMGKAND 109
          |::||||||||:|  :|  ||  |::||| |:|:||||||||:|||:||:||::|  :|||
Qy     65 DRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG--YG 122

Db    110 PDEMGIIPRLCNDLFARIDN-NNDKDVQYSVEVSYMEIYCERVKDLLNPSGGNLRVREH  168
          : :|||||||| ||::| |:  ::|||:|:::|::|||||||||||||||||||:|||||
Qy    123 KEH-GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181

Db    169 PLLGPVDDLTKMAVCSYHDICNLMDEGNKARTVAATNMNSTSSRSHAVFTIVLTQKRHC  228
          ||||||||||  ||:||:|||:|||||||||||||||||||||||||||||| ||||:|:
Qy    182 PSTGPVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241

Db    229 ADSNLDTEKHSKISLVDLAGSERANSTGAEGQRLKEGANINKSLTTLGLVISKLAEESTK 288
          |||:||||||:|||||||||||||| |||||||:||||:|||||||| ||||||||::
Qy    242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301

Db    289 KKKSNKGVIPYRDSVLTWLLRENLGGNSKTAMLAALSPADINFDETLSTLRYADRAKQIV 348
          ::||||||||||||||||| ||||||||||:::|||||||||||||::||||:|||||||
Qy    302 KQKKNQ-LVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK 360

Db    349 CQAVVNEDPNAKLIRELNEEVIKLRHILKDKG-----IDVTD--VQET--PGK--HKK-G 396
          ::|||||||:||| |||:|:||:||:| | |      |:|:|      :  | :
Qy    361 NHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGSGGGGSGGGGPVEESYPPDTPLEKQIV 420

Db    397 P-KLP-AHVH----EQL-EKLQESEKLMAEIGKTWEQKLIHTEEIRKQREEELRDMGLAC 449
          |  | | :||     |    |:|:||: :|| || |||::| ||:||||:| | ::|:
Qy    421 SIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEIHKEREAALELGIS- 479

Db    450 AEDGTTLGVFSPKKLPHLVNLNEDPLMSECLIYLKEGVSVGRPVAEHRPDILLSGEAI 509
          |::||||||||| ||| ||:|| :||||::||| :|: | |    :||||||:||||:|
Qy    480 IEKGF-VGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIRLNGSKI 538

Db    510 LELHCEFINEDGNVTLTMKPNASCYINGKQVTTPTVLHTGSRVILGEHHVFRYNDPQEAR 569
          | |:|| :|| :::| |  :| |:|:|| || : |||  | ::|:|  |:|:|| |::||
Qy    539 LKEHCTFENVDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEAR 598

Db    570 QSRHN 574
          ||
Qy    599 AERQE 603

RESULT 5
ID   Q9WVE5               PRELIMINARY;      PRT;  1816 AA.
AC   Q9WVE5;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KIF1B MAJOR BRAIN ISOFORM.
GN   KIF1B.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=BRAIN;
RA   CONFORTI L., BUCKMASTER A., TARLTON A., BROWN M.C., LYON M.F.,
RA   PERRY V.H., COLEMAN M.P.;
RT   "The major brain isoform of Kif1b lacks the putative mitochondria-
RT   binding domain.";
RL   Mamm. Genome 10:0(1999).
DR   EMBL; AF131865; AAD39438.1; -.
```

```
Db    167 VREHPILGPYVQDLSKLAVTSYADIADIADLMDCGNKARTVAATNMNETSSRSHAVFTIVFTQ 226
          ||||| ||||:||:||:|: :| |: :| |||||:||||||||||||||||||||||:: :|
Qy    178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237

Db    227 RSHDQLTGLDSEKVSKISLVNLAGSERADSSGARGMRLKEGANINKSLTTLGKVISALAD 286
          :|||:| :|:|::|::||||:|||||| |||||:| :|||||:|||||::||: |:||||
Qy    238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db    287 LQS-KKRKSDFIPYRDSVLTWLLKENLGGNSRTAMIAALSPADINYEETLSTLRYADRTK 345
          | |  ::  :|||||||||||||:|||||||:||||||:||||| ||||||||||||:|:
Qy    298 MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357

Db    346 QIRCNAVINEDPNARLIRELQEEVARLIRELLMAQGLSASALGGLKVEEGSPGGVLPAASS 405
          ||| |:|:||||||:|||||:||:||| |:|: :|||:||:| |||:|||||||  | ||
Qy    358 RIKNHAVVNEDPNARMIRELKEELAQLRSKL--QS-SGGG-GG--GAGGSGG--PVEES 408

Db    406 PPAPASPSSPPPHNGELEPSFSPSAEPQIGPEEAMERLQETEKIIAELNETWEEKLRKTE 465
          | | :  : ::: |::|  : :|: :||   :|: : |::|:  ::|::||:|||::||||
Qy    409 YP-PDTPLEKQIVSIQ-QPD-A-TVK-KMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTE 463

Db    466 ALRMEREALLAEMGSPGGWRTVGVFSPKKTPHLVNLNEDPLMSECLLYHIKDGVTRVGQV 525
          |:: |  : |: ||:||||| | ||||||:|||||||:|||:||||:|:||||:||::::
Qy    464 EIHKEREAALELGISIEKGFVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNV 523

Db    526 --DV--DIKLTGQFIREQHCLFRSIPQPDGEVMVTLEPCEGAETYVNGKLVTEPLVLKSG 581
          :|| |:|||||| |: ::| | :|||  |:||:| ::| || |||  |: : |
Qy    524 NQDTQAEIRLNGSKILKEHCTFENV---D-NV-VTIVPNEKAAVMVNGVRIDKPTRLRSG 578

Db    582 NRIVMGKNHVFRFNHPEQARLERE 605
          ||:|:|:||:||:||||:|:  |
Qy    579 YRIILGDFHIFRFNHPEEARAERQ 602

RESULT 4
ID   Q18778               PRELIMINARY;      PRT;  1584 AA.
AC   Q18778;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   C. ELEGANS UNC-104 KINESIN-LIKE PROTEIN (PIR:JN0114).
GN   UNC-104.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   DU Z.;
RP   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   WATERSTON R.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50135; AAA93453.1; -.
HSSP; P17119; 3KAR.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PFAM; PF00169; PH; 1.
```

RESULT 6

```
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE   1816 AA;  204154 MW;  D94A9653 CRC32;

Query Match             34.3%;  Score 1858;  DB 11;  Length 1816;
Best Local Similarity   52.6%;  Pred. No. 0.00e+00;
Matches  346;  Conservative  133;  Mismatches  110;  Indels  69;  Gaps  39;

Db          1 MSGAS-VKVAVRVRPFNSRETSKESKCIIQMGQNSTSIINP-----K--NP-K---EAPK 48
              :::: ::: ||||||||||:  : : :::   ::  ::  :     :   :: :   ::||
Qy          1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK 60

Db         49 SFSFDYSYWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNVCIFAYGQTGAGKSYTMM 108
              :.|| |||| |: : : ::: : :|||| | |:|| |  |:|||| |||||||||:|| ||
Qy         61 AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db        109 GKQEESQAVIIPQLCEELFEKIND-NCNEEMSYSVEVSYMEIYCERVRDLLNPKNKGNLR 167
              | | | ::   :       : |:| :| |: :|:::||:|| |:|||||||| ||||:
Qy        120 GYGKE-HGVI-PRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK 177

Db        168 VREHPLLGPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQ 227
              |||| ||||||||:||  | : : ||:||||||||||||||||||||||||||||::|||
Qy        178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237

Db        228 KKQDPETNLSTEVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE 287
              | :: ||  :||| |||||||||||||:|||| | ||||| ||:|| :||:||| ||||
Qy        238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db        288 VDNCTSKSKKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRY 347
              : :  : :.||:   ||||||||||||::.||||| |||:|| ||||||:::|||||||
Qy        298 M---SS-GKQKKN-QLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRY 352

Db        348 ADRAKQIKCNAVINEDPNAKLVRELKEEVTRLKDLLRAQGLGDIDIDPLIDDYSGSGGK 407
              ||.||:||:: ||||||||::|||||||: :|:..| :: ||:  : :: :: :| | |
Qy        353 ADSAKRIKNHAVVNEDPNARMIRELKEEALQLRSKLQSSG-GG-G--GA-GGSGGP 404

Db        408 YLKDFQNNKHRYLLASENQRPGNFSTASMGSLTSSPSSCSLNSQVGLTSVTSIQERIMST 467
                :     :  |  |: ::  :    :    :: :  : :   |::  |: :||||:|
Qy        405 -V-E-ES----Y----P----P-D--TP----L-EKQIV-SIQ-Q--PDAT-VK-K-MSK 434

Db        468 PGGEEAIERLKESEKIIAELNETWEEKLRKTEAIRMEREALLAEMGVAIREDGGTLGVFS 527
                ||||||||||| ||: ||:|||||||||:|||: :|||| ::|: | :|| |||| |:
Qy        435 A---EIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALEELGISI-EKG-FVGPYH 489

Db        528 PKKTPHLVNLNEDPLMSECLLYYIKDGITKRVGQADAERRQDIVLSGAHIKEEHCLFRSER 587
                : ||||||::||| :|||:|| || |:|:|   :: | :||:|:|| :||||| || |
Qy        490 SKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIRLNGSKILKEHCTF--E- 546

Db        588 SNTGEVIVTLEPCERSETYVNGKRVAHPVQLRSGNRIIMGKNHVFRFNHPEQARAERE 645
               |    | |: |  :   |:||:||| |  |||| ||| | :|:|||||||| ||||:
Qy        547 -NVDNV-VTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEARAERQ 602
```

RESULT   6
ID   O88658;   PRELIMINARY;   PRT;   689 AA.
AC   O88658;
DT   01-NOV-1998 (TREMBLrel. 08, Created)
DT   01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE   KINESIN-LIKE PROTEIN KIF1B (FRAGMENT).
GN   KIF1B.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   FAIRE K., GRUBER D., BULINSKI J.C.;
RT   "Identification of kinesin-like molecules in myogenic cells.";
RL   Eur. J. Cell Biol. 0:0-0(1998).
DR   EMBL; AF083331; AAC33292.1; -.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
FT   NON_TER   689   689

SQ (RESULT 7 alignment)

```
SQ   SEQUENCE   689 AA;  77352 MW;  E43A21ED CRC32;

Query Match             34.1%;  Score 1849;  DB 11;  Length 689;
Best Local Similarity   51.2%;  Pred. No. 0.00e+00;
Matches  329;  Conservative  130;  Mismatches  128;  Indels  56;  Gaps  34;

Db          1 MSGAS-VKVAVRVRPFNSRETSKESKCIIQMGQNSTSIINP-----K--NP-K---EAPK 48
              :::: ::: ||||||||||:  : : :::   ::  ::  :     :   :: :   ::||
Qy          1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK 60

Db         49 SFSFDYSYWSHTSPEDPCFASQSRVYNDIGKEMLLHAFEGYNVCIFAYGQTGAGKSYTMM 108
              :.|| |||| |: : : ::: : :|||| | |:|| |  |:|||| |||||||||:|| ||
Qy         61 AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db        109 GKQEESQAGIIPTCCEELFEKIND-NCNEDMSYSVS-SYMEIYCERVRDLLNPKNKGNLR 166
              | | | ::   :       : |:| :| |: :|:::||:|| |:|||||||| ||||:
Qy        120 GYGKE-H-GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK 177

Db        167 VREHPLLGPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQ 226
              |||| ||||||||:||  | : : ||:||||||||||||||||||||||||||||::|||
Qy        178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237

Db        227 KKQDPETNLSTEKVTKISLVDLAGSERADSTGAKGTRLKEGRNINKSLTTLGKVISALAE 286
              | :: ||  :||| |||||||||||||:|||| | ||||| ||:|| :||:||| ||||
Qy        238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db        287 VSK-KKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTSLRYADR 345
              : : .||:   ||||||||||||::.||||| |||:|| ||||||:::||| :|||||:
Qy        298 MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLST--LRYADS 355

Db        346 AKQIKCNAVINEDPNAKLVRELKEEVTRLKDLLRAQGLGDLLRAQGLGDIIDNLKDFQNNKHRYLLASE 405
              ||:||:: ||||||||::|||||||: :|:..| :: ||: :
Qy        356 AKRIKNHAVVNEDPNARMIRELKEEALQLRSKLQSSG-GG--G--AG------GSG 402

Db        406 NQRPGNFSTASMGSLTSSPSSCSLNSQAGLTSVTSIQERIMSTPGGEERAIERLKESEKII 465
                :     :    :: :  : :   |::  |: :||||:| :
Qy        403 GPVEESYPPDT--PL-EKQIV-SIQ-QPD--A-T-VK-K-MSKA--EIVEQLNQSEKLY 448

Db        466 AELNETWEEKLRKTEAIRMEREALLAEMGVAIREDGGTLGVFSPKKTPQVVNLNEDPLMS 525
                ||:|||||||||:|||: :|||| ::|: | :|| |||| |: :|| | ||::|||
Qy        449 RDLNQTWEEKLAKTEEIHKEREAALEELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLA 506

Db        526 ECLLYYIKDGITKGFGQADAERRQDIVLSGAHTIKEEHCIFRSERNNTGEVIVTLEPCER 585
              |||:|| || |:|  | :: | :||:|:|| :|||| || || ::|| |||:|:|||:|
Qy        507 ECLVYNIKPGQTRV-GNVNQDTQAEIRLNGSKILKE-HCTF--E--NVDNV-VTIVPNEK 559

Db        586 SETYVNGKRVAHPVQLRSGTVSSWVKNHVFRFNHPEQARAERE 628
               |:||:||| |  |||| | : ::|:|:|||||||| ||||:
Qy        560 AAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEARAERQ 602
```

RESULT   7
ID   O01349;   PRELIMINARY;   PRT;   1921 AA.
AC   O01349;
DT   01-JUL-1997 (TREMBLrel. 04, Created)
DT   01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE   KINESIN-73.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97188425.
RA   LI H.P., LIU Z.M., NIRENBERG M.;
RT   "Kinesin-73 in the nervous system of Drosophila embryos.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:1086-1091(1997).
DR   EMBL; U81788; AAB50404.1; -.
DR   FLYBASE; FBgn0019968; Khc-73.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PROSITE; PS00845; CAP_GLY; 1.
DR   PFAM; PF01302; CAP_GLY; 1.

```
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE   1921 AA;  215047 MW;  BF55A112 CRC32;

Query Match             29.5%;  Score 1597;  DB 5;  Length 1921;
Best Local Similarity   59.6%;  Pred. No. 1.37e-300;
Matches  238;  Conservative  66;  Mismatches  84;  Indels  11;  Gaps  8;

Db      1  MASD-KIKVAVRVRPFNRREIELDTKCIVEMEKQQTILQNPP-PLEKIER--K---Q-PK   52
              |::  || | ||  | ||||| | ||| ||| :||     :  | ||      : ||
Qy      1  MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK   60

Db     53  TFAFDHCFYSLNPEDENFASQETVFDCVGRGILDNAFQGYNACIFAYGQTGSGKSYTMMG  112
             :||  | :: : ||| | || :|||| | :|||| ||| |||||||||||||||||||
Qy     61  AFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG  120

Db    113  TQESKGIIPRLCDQLFSAIAN-KSTPELMYKVEVSYMEIYNEKVHDLLDPKPNKQSLKVR  171
             |:|| ||:|::: | :       | ||||::  :|| |||||:|::|:|:: : :|||
Qy    121  YGKEHGVIPRICQDMFRRINELQKDKNLTCVEVSYLEIYNERVRDLLNP-STKGNLKVR   179

Db    172  EHNVMGPYVDGLSQLAVTSYQDINLMTEGNKSRTVAATMNAESSRSHAVFSVVLTQIL   231
             ||||||||||| || |||||:|: |||||||  |||||||||||||||||||: | :
Qy    180  EHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATMNETSSRSHAVFTLTLTQKW  239

Db    232  TDQATGVSGEKVSRMSLVDLAGSERAVKTGAVGDRLKEGSNINKSLTTLGLVISKLADQS  291
             |: ::  | ||||:|:||||||| |  | |    |||| |||||::| |||::| | |
Qy    240  HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMS  299

Db    292  NGKKSGNDKFVPYRDSVLTWLLKDNLGGNSRTVMVATISPSADNYEETLSTLRYADRAKR  351
             ||||| || |||||||||||||| | |||:| |::| |:| || |||||||||||: :|
Qy    300  SGKQK-KNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKR  358

Db    352  IVNHAVVNEDPNARIIRELRHEVETLRSMLKHATGSPVG  390
             | ||||||||||||:||||||||||||: :|: ::| |
Qy    359  IKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGG  397


RESULT  8
ID   Q15058       PRELIMINARY;     PRT;   1648 AA.
AC   Q15058;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KIAA0042 PROTEIN.
GN   KIAA0042.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96051398.
RA   NOMURA N., NAGASE T., MIYAJIMA N., SAZUKA T., TANAKA A., SATO S.,
RA   SEKI N., KAWARABAYASHI Y., ISHIKAWA K., TABATA S.;
RT   "Prediction of the coding sequences of unidentified human genes. II.
RT   The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 1:223-229(1994).
DR   EMBL; D26361; BAA05392.1; -.
DR   HSSP; P17119; 3KAR.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE   1648 AA;  186491 MW;  ECACF441 CRC32;

Query Match             24.1%;  Score 1308;  DB 4;  Length 1648;
Best Local Similarity   43.9%;  Pred. No. 9.45e-239;
Matches  241;  Conservative 126;  Mismatches 133;  Indels  49;  Gaps  28;

Db    401  FIYDVSFWSFDECHPHYASQTTVVEKLAAPLLERAFEGENTCLFAYGQTGSGKSYTMMGF  460
             | :| |:||| |             ||| |  |  ::|||||||||||||||:||||
```

```
Qy      62  FAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY  121
              :|| | |:||:   :   :: :  ||| :|:|::   : : :|||||||||||||::||
Db     461  SEEPGIIPRFCEDLFSQVARKQTQ-EVSYHIEMSFFEVVNEKIHDLLVCKDENGQRKQPL  519

Qy     122  GKEHGVIPRICQDMFRRINELQKDKNLTCVEVSYLEIYNERVRDLL---NPST--KGNL  176
              ::| :|:||::| :|:|  ||: :|  | ::: |   |: ||:||  |: |   ||||
Db     520  RVREHPVIGPYVEALSMNIVSSYADIQSWLELGNKQRATAATGMNDKSSRSHSVFTLVMT  579

Qy     177  KVREHFSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATMNETSSRSHAVFTLTLT  236
             ||||| |||||::|| || ::|| :|:|:| | |||||||||||||||||||||:|:|
Db     580  QTKTEFVEGEEHDHRITSRINLIDLAGSERCSTAHTNGDRLKEGVSINKSLLTLGKVISA  639

Qy     237  QKWHD-EETKMDTEKVA-KISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAA  294
             |: |   |||::|||:  |||| |||:||:  : : :  :|:|| ::|: |  ||::| |
Db     640  LSEQAN--QR-SVFIPYRESVLTWLLKESLGGNSKTAMIATISPAASNIEETLSTLRYAN  696

Qy     295  LADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYAD  354
             || ||: :| ||||||||||||||||||:||||| ||||::|||:: |||||||||||:
Db     697  QARLIVNIAKVNEDMNAKLIRELKAEIAKLKAA-QRNSRN---ID----P---ERYR---  742

Qy     355  SAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGAGGSGGPVEESYPPDTP  414
             : || ||:  : |    | ::| :||   : :  :              :    : | |
Db     743  LCRQ--EI----TSLR-M-K--L-HQ-QERD-M-AEMQRVWKEKFEQAE--K-RK--LQ  782

Qy     415  LEKQIVSIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALE  474
             ||| : |   ||| |::::|||:  :|:|||  |: :|::|:| | ||:||:|: ||
Db     783  E-TKELQRAGI-MFQMDNHLPNLVNLNEDPQLSEMLLYMIKEGTTTVGKYKPNSSHDIQL  840

Qy     475  ELGISIEKGFVGPYH-SKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIRL  533
             || : :|||:|||:   ||:|:|  : :| |  | ||: | | | | :|:|  | || |
Db     841  SGVLIADDHCTIKNFGCTVSIPVGEAKTYVNGKHILETTVLRHGDRVILGDHYFRFNH  900

Qy     534  NGSKILKEHCTFENVDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNH  593
             |||::||:| ||:|  || |:||:|   |:|:|| || |||||||||:|||| | |||||
Db     901  PVEVQKGR  909
             |:|: |||
Qy     594  PEEA-RAER  601


RESULT  9
ID   Q09997       PRELIMINARY;     PRT;    928 AA.
AC   Q09997;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PUTATIVE KINESIN-LIKE PROTEIN R144.1 IN CHROMOSOME III.
GN   R144.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   FAVELLO T.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO A FAMILY OF KINESIN-LIKE PROTEINS
CC       CHARACTERIZED BY THE PRESENCE OF A MECHANOCHEMICAL DOMAIN TETHERED
CC       TO DIFFERENT PROTEIN BINDING DOMAINS.
DR   EMBL; U23515; AAC46545.1; -.
DR   HSSP; P17119; 3KAR.
DR   WORMPEP; R144.1; CE02031.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Hypothetical protein; Motor protein; Microtubules; ATP-binding;
KW   Coiled coil.
FT   DOMAIN        1    350       MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT   DOMAIN      445    448       POLY-LYS.
NP_BIND     107    114       ATP (BY SIMILARITY).
SQ   SEQUENCE    928 AA;  103589 MW;  CABE5681 CRC32;

Query Match             23.8%;  Score 1291;  DB 5;  Length 928;
Best Local Similarity   49.8%;  Pred. No. 3.96e-235;
Matches  199;  Conservative  81;  Mismatches 113;  Indels  7;  Gaps  6;
```

US-09-235-416-1.rspt

```
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN-BRISTOL N2;
RA   WATERSTON R.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U41536; AAB52613.1; -.
DR   HSSP; P33176; 1BG2.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00225; kinesin; 2.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE  1576 AA;  177409 MW;  8001SD5C CRC32;

Query Match             20.7%;  Score 1124;  DB 5;  Length 1576;
Best Local Similarity   54.5%;
Matches  225;  Conservative  65;  Mismatches  78;  Indels  45;  Gaps  19;

Db        8 SAVKVAIRVRPFNKRELDKTKSVRIQKEQCVL-HHP-IEEK--NS-KT------FTF   55
              : :|||||:|:||| :|    ||::||||  |   |  ||   :: ||         |:|
Qy        5 GNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAFAF   64

Db       56 DHSFCSTDPHSYDFASQETVSYHLGSGVVENAFSGYNACIFAYGQTGSGKSYTMMGTPDQ  115
              |:|: | :::| ::   ||| ||:|:|||||||||||:::::|||||
Qy       65 DRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGYGKE  124

Db      116 PGIIPRVCNDIFTRIQE-TSNSSLSFKVEVSYMEIYNERVRDLLDPKKSSSKALKVREHKI  174
              |:|||:|:| |:|    | :| ||||||||:||||||:||||||||:: |||||
Qy      125 HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGN-LKVREHPS  183

Db      175 LGPMVDGLSILAVNSFEQISNLLEEGNKSRTVAATNMNAESSRSHAVFLIVTQTLHDLE  234
              |||::||||||:|:::| ||||:||||||:|||||||||:||||:||||||| |: || :
Qy      184 TGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDEE  243

Db      235 NGFSGEKVAKISLVDLAGSERAGKTGAVGKRLEEGGNINKNLVSIFLRNDLEKKIDFKFS  294
              ||:|:|||||||||||||||:| |||::||:|| |::| ||:|||||||::|| :|:|::
Qy      244 TKMDTEKVAKISLVDLAGSERATSTGAFGARLKEGAEINRSL-ST-----L---G-R--  290

Db      295 ADVYFYFINQKLDFRSLTTLGMVISALAERNSKKDKFIPYRDSVLTWLLKDSLGGNSRTVM  354
              || : : :   :| : :::|||||   ||| :: :|||||||||||||||||||||||
Qy      291 ----V-IA------V-IA--DM--SS-G-KQ--KNQLVPYRDSVLTWLLKDSLGGNSMTAM  332

Db      355 IATLSPAADNYEETLSTLRYADRAKKIVNHAIINEDPNARVIRELREEVETLR  407
              |||||||||||||||||||||:||  |||||||||||||||:|||:::: |||:
Qy      333 IAISPADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR  385

RESULT  11
ID   Q9Y496                 PRELIMINARY;      PRT;    702 AA.
AC   Q9Y496;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KINESIN FAMILY MEMBER PROTEIN KIF3A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE-RETINA;
RA   WANG S.Y., BOST-USINGER L., HOANG E., FRAZER K.A., REICHELT S.,
RA   BURNSIDE B.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF041853; AAC72294.1; -.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE  702 AA;  80337 MW;  35968AD6 CRC32;

Query Match             17.7%;  Score 958;  DB 4;  Length 702;
Best Local Similarity   40.8%;
Matches  193;  Conservative  107;  Mismatches  138;  Indels  35;  Gaps  27;

Db       14 NVKVVVRCRPLNERE--K-SMCY-K-QAV-SV-DEMRGTI-TVHKTDSSN-EPPKTFTFD   64
              |:|||||| ||:|| ||   :     ::  ||   :: :|| ::|:||
```

```
Db            1 MGKGDSIIVAVRVRPFNDREXTRNCKLVIEMPDEETTVIRDPKTNDEKRFTYDHSYW-SH   59
                 | |:| |||:||||||:|  : ||| ||:||:|   :|  : ||:  |  : |    |
Qy            1 MSSGGNIKVVVRVRPFNAREIDRGAKCIVRM-EGNQTILTPPGAEEKARKSGKTIMDGP   59

Db           60 DGFS-EKKNGYLEPTDPHYADQRRVFEDLGRGVLANAWAGYNCSLFAYGQTGSGKSYSIV  118
                 ::|: |: |   :|:||  |  ||:|:|    |  || ||||:|||||||||||||::|
Qy           60 KAFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db          119 GFKNNKGIVPIVCEELFKQIAD-NKKKNMQPEVFVSMMEIYCEKVRDLLSSTPPPKGGLK  177
                 | |:|: |: :|::  |  |:  :: : |   |:|:||||:||||:||| :  | |:|:||
Qy          120 GYGKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLN--PSTKGNLK  177

Db          178 VREHPKNGFYVENLTTVPVNSFKEIEAKIEEGTKSRTIAATQMNATSSRAHTIVKITFNQ  237
                 ||||| |:| |:| :| ||||:|| :| |:||| ||||||:||| :|:|||||:| |::|
Qy          178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ  237

Db          238 KSSKQAGGTSMKK-SEINLVDLAGSERQSAAGTEGDRLKEGIVINQSLTTLGRVIKALHD  296
                 | |:|  | :|:|  : |||||||||||:| || || ||:||:|:||||:|||||:||:|
Qy          238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGAFGARLKEGAEINRSLSTLGRVIAALAD  297

Db          297 SQAKSGKKTQIPYRDSVLTCLLKNALGGNSKTIMIAAISPADINFEETLSTLRFADRAK  356
                 | |:|: |   :||||| |||:| |||||:|  |:||||||||:||||||||||::|:|
Qy          298 MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK  357

Db          357 SIKTNAVVNENQTERALRELREENLRLQSOIQGGTAGDAS   396
                 ||  |:||| | |:::|||:||||||:| ::  :::
Qy          358 RIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGG  397

RESULT  10
ID   Q20888                 PRELIMINARY;      PRT;    1576 AA.
AC   Q20888;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   SIMILAR TO KINESIN-RELATED PROTEIN. NCBI GI: 1109842.
GN   F56E3.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC   Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN-BRISTOL N2;
RX   MEDLINE; 94150718.
RA   WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA   BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA   CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA   GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA   JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA   LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA   PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA   SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA   THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA   WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN-BRISTOL N2;
RA   GATTUNG S., WU X.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN-BRISTOL N2;
RA   WATERSTON R.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN-BRISTOL N2;
RA   WATERSTON R.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases
```

```
Qy    6 NIKVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAFAFD  65
          :  ||  | ||::  :   ||:|  :  |:|:   |::  |    |   |:| :  ||
Db   65 -T--VF---GPE-SKQLDVY-NLTARPIIDSVLEYNGTIFAYGQTGKTFTMEGVRAI  116
          |    |:     ||: .:   |::::  ||     ||||||||||||    |   |
Qy   66 RSYWSFDKNAPNYARQEDLFQDLGV-PLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-GK 123
          |   | :||||||    :   ||| ||||||||||||||          |:|:
Db  117 PELRGIIPNSFAHIFGHIAKAEGDTRFL--VRVSYLEIYNEEVRDLLGKDQTQRLAVKER 174
          |     |   |  |  |       | |    |||||||   |  ||| | |  |:|:
Qy  124 -E-HGVIPRICQDMFRRINELQKQKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181
           | : |||  |||||||||||||   :|:  | ||||||||||| ||||  | |||  |:|
Db  175 PDVGVYIKDLSAYVVNNADDMDRIMTLGHKNRSVGATNMNEHSSRSHAIFTITIECSEKG 234
          || | ||  ||:.| | :||  |:  | | | ||||||||||||||||  |:: |
Qy  182 PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241
           ||||   |||| |||   ||  |   | |||| ||||||||||| ||| | |||   |
Db  235 IDGNMHV-RMGKLHLVDLAGSEROAKTGATGQRLKEATKINLSLSTLGNVISALVD---G 290
          | ::  | | || |||||||||  :| | | |||||| || || : ||  ||||    |
Qy  242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301
          || |||||| |||||||||||||||  | ||| |||||  |:|||||| |:||  |  |
Db  291 K--STH-VPYRNSKLTRLLQDSLGGNSKTMMCANIGPADYVNDETISTLRYANRAKNIKN 347
          |    |  | |:| ||||||||| ||||:  ||  || | | :|| | |||| |:||||
Qy  302 KQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKN 361
          | |||| |||| |||  || ||||||| ||: | ||||| |:||| |||||||| |||||
Db  348 KARINEDPKDALLRQPQKEIEELKKKLEE-GEEISGSDISGSEEDDDEEGEVGEDGEKRK  406
          ||  |: |  |||    |:||||  | |   | | | |:| |  |  |    |:| |
Qy  362 HAVVNEDPNARMIRELKEELAQLRSKLQSSGGPVEESYPPDTPLEKQIVS  421
           |  |||| |: |||||||| ||| ||  | || ||| |   |:| :|:
Db  407 KRRIQIGKKKVSPDKMIE-M-QA-KIDEE-RKALETKLDMEEEERNKARAELE  455
          |||:||  |   || |:|  | || |:|||  | |||||| || |: ||:|| |
Qy  422 IQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALE  474
```

```
RESULT 12
ID   093478
AC   093478;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KINESIN LIKE PROTEIN 3
GN   XKLP3.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93246065.
RA   VERNOS I., HEASMAN J., WYLIE C.;
RT   "Multiple kinesin-like transcripts in Xenopus oocytes.";
RL   Dev. Biol. 157:232-239(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   LE BOT N., ANTONY C., WHITE J., KARSENTI E., VERNOS I.;
RT   "Role of XKlp3, a subunit of Xenopus kinesin II heterotrimeric
RT   complex, in membrane transport at the ER/Golgi interface.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ009839; CAA08879.1; -.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE 744 AA; 84873 MW; 5674E1D1 CRC32;
```

```
Query Match       16.7%;  Score 904;  DB 13;  Length 744;
Best Local Similarity 46.5%;  Pred. No. 2.57e-153;
Matches  155;  Conservative  72;  Mismatches  92;  Indels 14;  Gaps 9;

Db   53 LSKTF-TFVATTDSNSKQVELYVETFRPLVDSVLLGLNGTIFATGQTGTCKTYTMEGVRG 111
          ::  |  |  ||:|:|: |:||||| || : ||: |||||||:|||||  |||:||||
Qy   64 FDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-G 122
Db  112 DPEKRGVIPNSFEHIFTHISRSQNQYL---VRASYLEIYQEIRDLLSKDQSKRLELKE 168
          : | |||||  | ||  || |:|||||     |:||||||:| ||| |||:|  :|
Qy  123 K-E-HGVIPRICQDMFRRINELQKQKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVRE 180
```

```
Db  169 RPDIGVYVKDLSFVTKSVKEIEHVMNVGNGQNRSVGATMNEHSSRSHAIFMITI-ECSQ 227
          :  | ||  |||:   |::| | :|  :  |:|||||:|||||||||| |:| | ||:
Qy  181 HPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWH 240
Db  228 IGLDGENHIRVGKLNLVDLAGSEROTKTGAQGERLKEATKINLSLSALGNVISALVD--- 284
          |  |   ||  |::| |||||||:|:| | | |||||| ||||||:|||  |:||
Qy  241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSS 300
Db  285 GR--STH-IPYRDSKLTRLLQDSLGGNAKTVMVANIGPASYNVEETLTTLRYSNRAKNIK 341
          |    ||  ||| | |||| ||||| |||:  ||  |||:   |:|| |||| |:|||||
Qy  301 GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK 360
Db  342 NKPRVNEDPKDALLREFQEEIARLKAQLDKRVG 374
          | : ||||| ||||| |||||  ||:| |:
Qy  361 NHAVVNEDPNARMIREKEELAQLRSKLQSSGG 393
```

```
RESULT 13
ID   018389
AC   018389;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KINESIN-LIKE PROTEIN, KLP38B.
GN   KINESIN-LIKE PROTEIN, KLP38B.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97296456.
RA   OHKURA H., TOROK T., TICK G., HOHEISEL J., KISS I., GLOVER D.M.;
RT   "Mutation of a gene for a Drosophila kinesin-like protein, Klp38B,
RT   leads to failure of cytokinesis.";
RL   J. Cell Sci. 110:945-954(1997).
DR   EMBL; Y10667; CAA71675.1; -.
DR   HSSP; P17119; 3KAR.
DR   FLYBASE; FBgn0004374; neb.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE 1121 AA; 125194 MW; B0A065F6 CRC32;
```

```
Query Match       15.8%;  Score 857;  DB 5;  Length 1121;
Best Local Similarity 37.3%;  Pred. No. 1.71e-143;
Matches  243;  Conservative  147;  Mismatches 179;  Indels 83;  Gaps 42;

Db  115 VSEESNMIVAVRVRPLNALECTRGQVTNVGVHGNSNELTVQAGSSADAS-AG--VTH-- 169
          :  ::|| | ||||| |   | | ||:|| |:|:|||  | | ||:|:      |:|
Qy    1 MSGGGNIKVVVRVRPFNAREIDRG-AKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGP  59
Db  170 -FFSYDQVYYSCDPERKNFACQAKVFEGTARPLIDTAFEGYNACLFAYGQTGSGKSYSMM 228
           |: ||  :   : |:  ||   :    | | |:|| ||||| |:|||||||||||||||
Qy   60 KAFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119
Db  229 GIEALDDAALDGGPPHDEAGIIPRFCHELFRRIEAVKSQQQLQVEVEVSYFEIYNEKIHD 288
          |        |    ||  :  ||||  |:||| || |||| ||||||||||||||:: |
Qy  120 GY-G--K---E----H--GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRD 166
Db  289 LLSVQHAAAATGESTPIQQQQQQQRPALKVREHPIFGPYVVDLSAHSVDSYSALRNWLAV 348
          ||      |||||:| |             |||| ||| |:||:|  ||| | | |:|
Qy  167 LL---NPST-KGN-----------LKVREHPSTGPVYEDLAKLVVRSFQEIENLMDE 208
Db  349 GNSQRATASTAMNDKSSRSHSIFNIVLNLTDLSSDDGLSSDTDSSTASSLRQTRRSKISL 408
          || |:  | |:|:| |||||:::: |  |  |  |:   |:   | |||| : : ||||
Qy  209 GNKARTVAATNMNETSSRSHAVFT--LTLTQ-KWHD----E-ETKMDTE-K-V--AKISL 256
Db  409 VDLAGSERISVGGSNGERIREGVSINKSLLTLGKVIAALADSRKASANGPLGSGTPSTFV 468
          ||||||||:|   |:|  |:||  |||| ||||:||||||||      | :   | : |
Qy  257 VDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD--MSSGK-Q--K-K-NQLV 309
```

```
Db  469  PYRESVLTWLLRENLGGNSKTVMLATISPASIHADETLATLRYACKARSIVNRVKVNESP  528
         ||||:|| |||||  |:|: ||||||||||| ::|  ||  :|||
Qy  310  PYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLSTLRYADSAKRIKNHAVVNEDP   369

Db  529  HDKIIRDLRAEVDRLKS-LRNEYERQRRL--SGNS-NNPVPRKIIIET---SVDETE--V  579
         : :|| : : |: |    |||  : : :: ::|
Qy  370  NARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPDTPLEKQIVSIQQPDATV    429

Db  580  EAL-RQQLAER--ERE-LSRA-QKSWMEKLKEAEDQRKSELRVLKRRGLALE--LTAE-Q  631
         :|| ||: : |      |:|  ||  :: :|||    : | || :  : |||
Qy  430  KKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALEELGISIEKGFVGPYH  489

Db  632  -KQ-ACLVNLTADPILSGTLFYLLPQGLVRIGRGRLPGGSSSSQPDIVLDGPLVALQHCS  689
          :: | |||| :||:|  ||  :|  ||:| |    :  : :|||  ::|
Qy  490  SKEMPHLVNLSDDPLLAECLVYNIKPGQTRVG-N-V---NQDTQAEIRLNGSKILKEHCT  544

Db  690  IEHERGGKLYVIPGSEDFETYVNGELLKDRRQLFHGDRLVIGGSHYFRISNP          741
         |||      |::| |:||| |:||  :|      || |:::|||  :|| :| ::||
Qy  545  FENV-DNVVTIVP-NEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHP          594
```

RESULT 14
ID   O02030;
AC   O02030;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KINESIN-LIKE PROTEIN.
GN   KLP38B.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97375669.
RA   ALPHEY L., PARKER L., HAWCROFT G., GUO Y., KAISER K., MORGAN G.;
RT   "KLP38B: a mitotic kinesin-related protein that binds PP1.";
RL   J. Cell Biol. 138:395-409(1997).
DR   EMBL; X99617; CAA67928.1; -.
DR   HSSP; P17119; 3KAR.
DR   FLYBASE; FBgn0004374; neb.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE   1121 AA;   125235 MW;   AE508F66 CRC32;

Query Match   15.8%;   Score 857; DB 5; Length 1121;
Best Local Similarity   37.3%;   Pred. No. 1.71e-143;
Matches 243; Conservative 147; Mismatches 179; Indels 83; Gaps 42;

```
Db  115  VSEESNMIVAVRVRPLNALECTRGQVTNVVQVHGNSNELTVQAGSSADAS-AG--VTH--  169
          :: |:||.||||||.|| :: | :|: |:  : ||. ::|  ||::||  |    ::
Qy    1  MSGGGNIKVVVRVRPFNAREIDRG-AKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGP   59

Db  170  -FFSYDQVYYSCDPERKNFACQAKVFEGTARPLIDTAFEGYNACLFAYGQTGSGKSYSMM  228
          :|: | ||:|: ::  |:  |   |::|||||:| ||.||| |:|||||||||||||||
Qy   60  KAFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db  229  GIEALDDAALDGGPPHDEAGIIPRFCHELFRRIEAVKSQQQLQVEVEVSYFEIYNEKIHD  288
         |  | :        :     :||| | ::||||.: :     ||:  |: |||||:: |
Qy  120  GY-G--K----E----H----GVIPRICQDMFRRINELQKDRNLTCVEVSLEIYNERVRD  166

Db  289  LLSVQHAAAATGESTPIQQQQORPALKVREHPIFGPYVVDLSAHSVDSYSALRNWLAV    348
         ||  :          :          ||||| :||||| || | |:| |: :
Qy  167  LL----NPST-KGN-------LKVREHPSTGPYVEDLAKLVVRSFQEIENLMDE       208

Db  349  GNSQRATASTAMNDKSSRSHSIFNIVLNLTDLSSDDGLSSDTDSSTASSLRQTRRSKISL  408
         || : .:.: ||:: |||||::| : |:|     :  ::    :  |:|||||||||||
Qy  209  GNKARTVAATMNNETSSRSHAVFT--LTLTQ-KWHD----E-ETKMDTE-K-V--AKISL  256
```

```
Db  409  VDLAGSERISVSGSNGERIREGVSINKSLLTLGKVIAALADSRKASANGPLGSGTPSTFV  468
         |||||||| :|:|:.|.|:  |  |:||| ||| |||||||   ::|| |    |  : |
Qy  257  VDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD--MSSGK-Q--K-NQLV   309

Db  469  PYRESVLTWLLRENLGGNSKTVMLATISPASIHADETLATLRYACKARSIVNRVKVNESP  528
         ||| :||||||::.|||||  . :|:|||:.|. .|||:|||| :.|||.||:: |:| |
Qy  310  PYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLSTLRYADSAKRIKNHAVVNEDP   369

Db  529  HDKIIRDLRAEVDRLKS-LRNEYERQRRL--SGNS-NNPVPRKIIIET---SVDETE--V  579
Qy  370  NARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPDTPLEKQIVSIQQPDATV    429

Db  580  EAL-RQQLAER--ERE-LSRA-QKSWMEKLKEAEDQRKSELRVLKRRGLALE--LTAE-Q  631
Qy  430  KKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALEELGISIEKGFVGPYH  489

Db  632  -KQ-ACLVNLTADPILSGTLFYLLPQGLVRIGRGRLPGGSSSSQPDIVLDGPLVALQHCS  689
Qy  490  SKEMPHLVNLSDDPLLAECLVYNIKPGQTRVG-N-V---NQDTQAEIRLNGSKILKEHCT  544

Db  690  IEHERGGKLYVIPGSEDFETYVNGELLKDRRQLFHGDRLVIGGSHYFRISNP          741
Qy  545  FENV-DNVVTIVP-NEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHP          594
```

RESULT 15
ID   O16866;
AC   O16866;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   KINESIN LIKE PROTEIN AT 38B.
GN   KLP38B.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   RUDEN D.M., CUI W., SOLLARS V., ALTERMAN M.;
RL   Dev. Biol. 0:0-0(1997).
DR   EMBL; AF022650; AAB81511.1; -.
DR   HSSP; P17119; 3KAR.
DR   FLYBASE; FBgn0004374; neb.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR   PFAM; PF00498; FHA; 1.
DR   PFAM; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil.
SQ   SEQUENCE   1121 AA;   125236 MW;   A43A76BF CRC32;

Query Match   15.8%;   Score 854; DB 5; Length 1121;
Best Local Similarity   37.3%;   Pred. No. 7.23e-143;
Matches 243; Conservative 146; Mismatches 180; Indels 83; Gaps 41;

```
Db  115  VSEESNMIVAVRVRPLNALECTRGQVTNVVQVHGNSNELTVQAGSSADAS-AG--VTH--  169
Qy    1  MSGGGNIKVVVRVRPFNAREIDRG-AKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGP   59

Db  170  -FFSYDQVYYSCDPERKNFACQAKVFEGTARPLIDTAFEGYNACLFAYGQTGSGKSYSMM  228
Qy   60  KAFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM  119

Db  229  GIEALDDAALDGGPPHDEAGIIPRFCHELFRRIEAVKSQQQLQVEVEVSYFEIYNEKIHD  288
Qy  120  GY-G--K----E----H----GVIPRICQDMFRRINELQKDRNLTCVEVSLEIYNERVRD  166

Db  289  LLSVQHAAAATGESTPIQQQQQRPALKVREHPIFGPYVVDLSAHSVDSYSALRNWLAV    348
Qy  167  LL----NPST-KGN-------LKVREHPSTGPYVEDLAKLVVRSFQEIENLMDE       208

Db  349  GNSQRATASTAMNDKSSRSHSIFNIVLNLTDLSSDDGLSSDTDSSTASSLRQTRRSKISL  408
Qy  209  GNKARTVAATMNNETSSRSHAVFT--LTLTQ-KWHD----E-ETKMDTE-K-V--AKISL  256
```

US-09-235-416-1.rspt

Mon Aug 21 15:58:39 2000

```
            ||  |: |:|  ||: ||||||::|  |  ||:      |  :  ::  :  ::::: :  :|||||
Qy   209 GNKARTVAATNMNETSSRSHAVFT--LTLTQ-KWHD----E-ETKMDTE-K-V--AKISL 256
            ||||||| .:||: | |::|| ||::|| ||:|||||
Db   409 VDLAGSERISVSGSNGERIREGVSINKSLLTLGKVIAALADSRKAIANGPLGSGTPSTFV 468
            |||||||::|:: | |::|  |||:||  ::|    ::|| :  :|
Qy   257 VDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD----MSSGKQ-K-K-NQLV 309
            |||:||||||::|||| | :|:|||  |::|||||  |::|:||||  |:| |
Db   469 PYRESVLTWLLRENLGGNSKTVMLATISPASIHADETLATLRYACKARSIVNRVKVNESP 528
            |||:|||||||:::||||:| |::|||||  ||:|||||  |::| |||  |||  |
Qy   310 PYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKNHAVVNEDP 369
            ||:|::|| |: ::|| | |:   |   |  ::: |     :| ::::: |
Db   529 HDKIIRDLRAEVDRLKS-LRNEYERQRRL--SGNS-NNPVPRKIIIET---SVDETE--V 579
            : ::|||::| |: :|:| |::   ||  |:| :: :: : :|  |:::|    :|
Qy   370 NARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPDTPLEKQIVSIQQPDATV 429
            : ::|   |||:   ::|:  | :  | |  ||   |: ||: |: | :
Db   580 EAL-RQQLAER--ERE-LSRA-QKSWMEKLKEAEDQRKSELRVLKRRGLALE--LTAE-Q 631
            :  ::| |:  | |: |  :||  || :::||  ::||  :  ::::|
Qy   430 KKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALEELGISIEKGFVGPYH 489
            |:: :|||: |:  | |  :|: | |  :|||  ||  |:| ||:||  :|::
Db   632 -KQ-ACLVNLTADPILSGTLFYLLPQGLVRIGRGRLPGGSSSSQPDIVLDGPLVALQHCS 689
            |: : ||||  ||::|   |  |:|   | |:    :|::| |||: ::|:
Qy   490 SKEMPHLVNLSDDPLLAECLVYNIKPGQTRVG-N-V---NQDTQAEIRLNGSKILKEHCT 544
            |:  |:   | ||| :| || ||:||  |: :|:| |||:||:
Db   690 IEHERGGKLYVIPGSEDFETVVNGELLKDRRQLFHGDRLVIGGSHYFRISNP 741
            :|:  :| | ||:  ||  ::|: |  :|  :  | :|::::|
Qy   545 FENV-DNVTIVP-NEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHP 594
```

Search completed: Mon Aug 21 15:42:14 2000
Job time : 66 secs.

THIS PAGE BLANK (USPTO)

```
*********************************************

   M P S R C H    (TM)

*********************************************
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 21 15:42:31 2000:   MasPar time 15.62 Seconds
                                               724.633 Million cell updates/sec

Tabular output not generated.

Title: >US-09-235-416-1
Description: (1-784) from US09235416A.pep
Perfect Score: 5422
Sequence: 1 MSGGGNIKVVVRVRPFNARE.........ELRQQAQMEEALKTAKQEF 784

Scoring table: PAM 150
                  Gap 11

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 0%
                  Listing first 45 summaries

Database: a-issued
                  1:5A_COMB 2:5B_COMB 3:6_COMB 4:PCT_COMB 5:backfiles1

Statistics: Mean 35.470; Variance 183.982; scale 0.193

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2037	37.6	1103	3	US-09-162-	Sequence 1, Applicatio	1.07e-153
2	953	17.6	955	4	PCT-US94-0	Sequence 1, Applicatio	7.47e-65
3	953	17.6	955	2	US-08-428-	Sequence 3, Applicatio	7.47e-65
4	953	17.6	955	1	US-08-006-	Sequence 1, Applicatio	7.47e-65
5	953	17.6	955	1	US-08-282-	Sequence 2, Applicatio	7.47e-65
6	707	13.0	411	2	US-08-713-	Sequence 4, Applicatio	5.23e-45
7	707	13.0	441	1	US-08-713-	Sequence 4, Applicatio	5.23e-45
8	196	3.6	1612	4	PCT-US94-0	Sequence 48, Applicati	2.01e-05
9	196	3.6	1612	3	US-08-545-	Sequence 48, Applicati	2.01e-05
10	120	2.2	1354	2	US-08-685-	Sequence 6, Applicatio	2.95e+00
11	101	1.9	594	2	US-08-785-	Sequence 30, Applicati	4.45e+01
12	95	1.8	240	4	PCT-US94-1	Sequence 30, Applicati	1.01e+02
13	95	1.8	240	1	US-08-362-	Sequence 30, Applicati	1.01e+02
14	95	1.8	240	3	US-08-333-	Sequence 30, Applicati	1.01e+02
15	98	1.8	274	2	US-08-860-	Sequence 10, Applicati	6.74e+01
16	97	1.8	280	2	US-08-403-	Sequence 10, Applicati	7.72e+01
17	100	1.8	371	2	US-08-442-	Sequence 76, Applicati	5.12e+01
18	95	1.8	414	2	US-08-132-	Sequence 2, Applicatio	1.01e+02
19	95	1.8	414	4	PCT-US91-0	Sequence 2, Applicatio	1.01e+02
20	95	1.8	414	4	PCT-US94-0	Sequence 6, Applicatio	1.01e+02
21	95	1.8	414	5	5221620-4	atent No. 5221620	1.01e+02
22	95	1.8	414	1	US-08-395-	Sequence 2, Applicatio	1.01e+02
23	100	1.8	425	2	US-08-986-	Sequence 2, Applicatio	5.12e+01
24	100	1.8	442	2	US-08-687-	Sequence 6, Applicatio	5.12e+01
25	95	1.8	442	5	5221620-2	atent No. 5221620	1.01e+02
26	96	1.8	470	2	US-08-933-	Sequence 6, Applicatio	8.85e+01
27	96	1.8	491	1	US-08-489-	Sequence 3, Applicatio	8.85e+01
28	96	1.8	491	2	US-08-993-	Sequence 3, Applicatio	8.85e+01
29	97	1.8	517	2	US-08-796-	Sequence 8, Applicatio	7.72e+01
30	97	1.8	566	2	US-08-533-	Sequence 8, Applicatio	7.72e+01
31	97	1.8	566	2	US-08-511-	Sequence 2, Applicatio	7.72e+01
32	95	1.8	610	3	US-08-525-	Sequence 8, Applicatio	1.01e+02
33	97	1.8	751	3	US-08-946-	Sequence 6, Applicatio	7.72e+01
34	96	1.8	760	1	US-08-195-	Sequence 2, Applicatio	8.85e+01
35	100	1.8	924	1	US-08-656-	Sequence 28, Applicati	5.12e+01
36	100	1.8	924	1	US-08-481-	Sequence 28, Applicati	5.12e+01
37	100	1.8	924	1	US-08-485-	Sequence 28, Applicati	5.12e+01
38	100	1.8	924	2	US-08-487-	Sequence 28, Applicati	5.12e+01
39	98	1.8	1093	4	PCT-US94-0	Sequence 55, Applicati	6.74e+01
40	98	1.8	1093	3	PCT-US94-	Sequence 55, Applicati	6.74e+01
41	94	1.7	312	1	US-08-414-	Sequence 2, Applicatio	1.16e+02
42	93	1.7	368	2	US-08-896-	Sequence 4, Applicatio	1.33e+02
43	94	1.7	802	2	US-08-007-	Sequence 4, Applicatio	1.16e+02
44	92	1.7	876	4	PCT-US95-0	Sequence 34, Applicati	1.52e+02
45	92	1.7	876	4	PCT-US95-0	Sequence 32, Applicati	1.52e+02

ALIGNMENTS

```
RESULT  1
ID   US-09-162-373-1          STANDARD;     PRT;    1103 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 1, Application US/09162373
XX
CC   Sequence 1, Application US/09162373
CC   Patent No. 6013454
CC   GENERAL INFORMATION:
CC   APPLICANT: Tang, Y., Tom
CC   APPLICANT: Corley, Neil C.
CC   APPLICANT: Guegler, Karl J.
CC   APPLICANT: Patterson, Chandra
CC   TITLE OF INVENTION: KINESIN-LIKE MOTOR PROTEIN
CC   FILE REFERENCE: PF-0593 US
CC   CURRENT APPLICATION NUMBER: US/09/162,373
CC   CURRENT FILING DATE: 1998-09-28
CC   NUMBER OF SEQ ID NOS: 2
CC   SOFTWARE: PERL Program
CC   SEQ ID NO 1
CC   LENGTH: 1103
CC   TYPE: PRT
CC   ORGANISM: Homo sapiens
CC   FEATURE: -
CC   OTHER INFORMATION: 1281811
SQ   SEQUENCE 1103 AA; 122860 MW; 5804959 CN;

Query Match            37.6%;  Score 2037;  DB 3;  Length 1103;
Best Local Similarity 54.2%;            Pred. No. 1.07e-153;
Matches  339;  Conservative 127;  Mismatches 118;  Indels 42;  Gaps 28;

Db    1 MAGAS-VKVAVRVRPFNARETSQDAKCVVSMQGNTTSIINP-----K-Q-S-K----DAPK 48
          |:|:: :|| |||||||||  : |||:| |:|| | ::|||    | :|  | |:|
Qy    1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPK 60

Db   49 SFTFDYSYWSHTSTEDPQFASQQQVYRDIGEEMLLHAFEGYNVCIFAYGQTGAGKSYTMM 108
         :|:|: :|| :  ::: : |:| |||||:: :|||||  |||||||:|||||||||
Qy   61 AFAFDRSYWSF-DKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db  109 GRQEPGQQGIVPQLCEDLFSRVSENQSAQ-LSYSVEVSYMEIYCERVRDLLNPKSRGSLR 167
         |:||::|:|:|:|| |:|| ||||||| : ||||:||| | |||||||||||| ::|:|
Qy  120 G-YGK-EHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK 177

Db  168 VREHPILGPYVQDLSKLAVTSYADIADLMDCGNKARTVAATNMNETSSRSHAVFTIVFTQ 227
```

```
CC    SEQUENCE CHARACTERISTICS:
CC       LENGTH: 955 amino acids
CC       TYPE: amino acid
CC       TOPOLOGY: linear
CC       MOLECULE TYPE: protein
SQ       SEQUENCE  955 AA;  106167 MW;  3654125 CN;

      Query Match          17.6%;  Score 953;  DB 4;  Length 955;
      Best Local Similarity 43.5%;  Pred. No. 7.47e-65;
      Matches  177;  Conservative  90;  Mismatches 111;  Indels 29;  Gaps 18;

Db     13  VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVTVKVLGGSNNSGAAESMGTARRVAQD  72
           : ::||||||:|  |  :| :  ::  :|   ::  :  |::
Qy      7  IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPGAEEKARKSGK-TIMDGPKA  61

Db     73  FQFDHVFWSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM 132
           | : : : ::|  | :: ::   |  ::|   :|||::|||||||||||:::
Qy     62  FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db    133  MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEVTNERVSDLLGKRK 192
           || : |  |   : |:||||  :| | :::|  : ||:|:::||||||||:
Qy    119  MG-----Y-GKEHGVIPRICQDMPRRINELQKDKNLTCVEVSYLEIYNERVRDLL-NP- 170

Db    193  KGVKGGGEVYVDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
            ::| |  | : |:||||||| :| |: :  ::|:   |  ::|::|| |
Qy    171  -STKG--NL-KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db    252  SSRSHAIIMLLRERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSQVEGQQFKEAT 311
           ||||||  : | :   : ::| :   ::::||||||||||||||:
Qy    224  SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db    312  HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
           :|| || |||||  :||||:: |:|: || | |||| | :|||||||||| | :|||:||
Qy    280  EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db    372  SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
           : :|:||||||||||: : :||::| |||||:|||:|||:|| || :
Qy    339  ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385


RESULT  3
ID   US-08-428-414A-3      STANDARD;      PRT;      955 AA.
XX
AC   xxxxxx
XX
DT
XX
DE
XX
CC   Sequence 3, Application US/08428414A
CC   Patent No. 5912166
CC   GENERAL INFORMATION:
CC      APPLICANT:  Reed, Steven G.
CC      TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF
CC      TITLE OF INVENTION:  LEISHMANIASIS
CC      NUMBER OF SEQUENCES:  5
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  SEED and BERRY
CC         STREET:  6300 Columbia Center, 701 Fifth Avenue
CC         CITY:  Seattle
CC         STATE:  Washington
CC         COUNTRY:  USA
CC         ZIP:  98104-7092
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/428,414A
CC         FILING DATE:  21-APR-1995
CC         CLASSIFICATION:  436
CC      ATTORNEY/AGENT INFORMATION:
```

```
Qy    178  VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237
           ||||| :  : : ::| |  :  :|        :  ||||||||||:: ||||
Db    228  RCHDQLTGLDSEKVSKISLVDLAGSERADSSGARGMGLKEGANINKSLTTLGKVISALAD 287
           : |:  : :||||||||||||||||| :|| : :::||||||:||||||||||:||:|||
Qy    238  KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db    288  MQS-KKRKSDFIPYRDSVLTWLLKENLGGNSRTAMIAALSPADINYEETLSTLRYADRTK 346
           | | ||:|::  ::|||||||||||::||||||:|||||:||||||:||||||||||:||
Qy    298  MSSGKGKKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357

Db    347  QIRCNAIINEDPNARLIRELQEEVARLRELLMAQGLSASALEGLKTEEGSVRGALPAVSS 406
            |: :|::||||||||:|||::|::||: :||| :| || |: ::  ::  ::
Qy    358  RIKNHAVVNEDPNARMIRELKEELAQLRSKL--QS-SGGG-GG--GAGGSG-GPVEE-SY 409

Db    407  PPAPVSPSSPTTHNGELEPSFSPNTESQIGPEEAMERLQETEKIIAELNETWEEKLRKTE 466
           ||  : |:  :  : ::|::   :: :  :: ||  :::|: :|::: ::|:||||||:||
Qy    410  PPD--TPLE--KQIVSIQQP-DA-TVKKMSKAEIVBQLNQSEKLYRDLNQTWEEKLAKTE 463

Db    467  ALRMEREALLAEMGVAVREDGGTVGVFSPKKTPHLVNLNEDPLMSECLLYHIKDGVTRVG 526
            :::|||| | : | : ::|  :||||  :||||||||::|||:: ||:|:|||| ||||
Qy    464  EIHKEREAALELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVG 521

Db    527  QV--DM--DIKLTGQFIREQHCLFRSIPQPDGEVVTLEPCEGAETYVNGKLVTEPLVLK 582
           :|  ::  :|:|  :||::||||:|:|  | : ::|   :: | :::|| ||:: | |||
Qy    522  NVNQPTQAEIRLNGSKILKEHCTFENV---D-NVV-TIVPNEKAAVMVNGVRIDKPTRLR 576

Db    583  SGNRIVMGKNHVFRFNHPEQARLERE 608
           || ||::| :| :||||||::||:|:
Qy    577  SGYRIILGDFHIFRFNHPEEARAERQ 602


RESULT  2
ID   PC-US94-00324-1      STANDARD;      PRT;      955 AA.
XX
AC   xxxxxx
XX
DT
XX
DE
XX
CC   Sequence 1, Application PC/TUS9400324
CC   GENERAL INFORMATION:
CC      APPLICANT:  Reed, Steven
CC      TITLE OF INVENTION:  Diagnosis of Leishmaniasis
CC      NUMBER OF SEQUENCES:  3
CC      CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Immunex Corporation
CC         STREET:  51 University Street
CC         CITY:  Seattle
CC         STATE:  Washington
CC         COUNTRY:  USA
CC         ZIP:  98101
CC      COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  Apple Macintosh
CC         OPERATING SYSTEM:  Apple System 7.1
CC         SOFTWARE:  Microsoft Word, version 5.1a
CC      CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US94/00324
CC         FILING DATE:
CC         CLASSIFICATION:
CC      PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/006,676
CC         FILING DATE:  15-JAN-1993
CC      ATTORNEY/AGENT INFORMATION:
CC         NAME:  Perkins, Patricia Anne
CC         REGISTRATION NUMBER:  34,693
CC         REFERENCE/DOCKET NUMBER:  5004-WO
CC      TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (206) 587-0430
CC         TELEFAX:  (206) 233-0644
CC   INFORMATION FOR SEQ ID NO:  1:
```

```
CC   NAME: Kadlecek, Ann T.
CC   REGISTRATION NUMBER: 39,244
CC   REFERENCE/DOCKET NUMBER: 210121.407
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (206) 622-4900
CC   TELEFAX: (206) 682-6031
CC   TELEX: 3723836 SEEDANDBERRY
CC   INFORMATION FOR SEQ ID NO: 3:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 955 amino acids
CC   TYPE: amino acid
CC   STRANDEDNESS:
CC   TOPOLOGY: linear
SQ   SEQUENCE   955 AA; 106152 MW; 3653591 CN;

Query Match             17.6%;  Score 953; DB 2; Length 955;
Best Local Similarity   43.5%;  Pred. No. 7.47e-65;
Matches   177;  Conservative   90;  Mismatches 111;  Indels 29;  Gaps 18;

Db    13  VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVVTVKVLGGSSNNSGAAESMGTARRVAQD 72
          ::||  |||||:| || :|          |:::          || ||  :     :|   :
Qy     7  IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA 61

Db    73  FQPDHVFWSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM 132
          |  |: :|   :  ||      ||  :  ||:| ||||    :| :|||||||||||:|:|
Qy    62  FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db   133  MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEVYNERVSDLLGKRK 192
          ||     :     ||  |||  ::|  |::: : || :| ::|:||||||||| |||   
Qy   119  MG-----Y-GKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLL-NP- 170

Db   193  KGVKGGGEEVIYDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
          :|| | :|  :|||||||| |  ||: |||:  |:        : || ||||   ::||: 
Qy   171  -STKG---NL--KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db   252  SSRSHAIIMLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSVEGQQFKEAT 311
          |||||| :  | |  |   |||| | || :: | ::|||||||||  : :| :    | 
Qy   224  SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db   312  HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
          ||| ||:||||||  |||| :|  |: ||::|| || | ||||||||||| | ||||:|||
Qy   280  EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db   372  SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
          : :|||||||||||| :|| |  ||:| |||  | |||::: ::|:|
Qy   339  ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385


RESULT 4
ID   US-08-006-676B-1      STANDARD;     PRT;    955 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 1, Application US/08006676B
XX
CC   Sequence 1, Application US/08006676B
CC   Patent No. 5411865
CC   GENERAL INFORMATION:
CC   APPLICANT: Reed, Steven
CC   TITLE OF INVENTION: Diagnosis of Leishmaniasis
CC   NUMBER OF SEQUENCES: 3
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Jeffrey B. Oster
CC   STREET: 8339 SE 57th Street
CC   CITY: Mercer Island
CC   STATE: Washington
CC   COUNTRY: USA
CC   ZIP: 98040-4906
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatable
```

```
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: WORD for Windows
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/006,676B
CC   FILING DATE: 15-JAN-1993
CC   CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Oster, Jeffrey B.
CC   REGISTRATION NUMBER: 32,585
CC   REFERENCE/DOCKET NUMBER: REED-4
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (206) 232 7845
CC   TELEFAX: (206) 236 0205
CC   INFORMATION FOR SEQ ID NO: 1:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 955 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE   955 AA; 106167 MW; 3654125 CN;

Query Match             17.6%;  Score 953; DB 1; Length 955;
Best Local Similarity   43.5%;  Pred. No. 7.47e-65;
Matches   177;  Conservative   90;  Mismatches 111;  Indels 29;  Gaps 18;

Db    13  VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVVTVKVLGGSSNNSGAAESMGTARRVAQD 72
          ::||  |||||:| || :|          |:::          || ||  :     :|   :
Qy     7  IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA 61

Db    73  FQPDHVFWSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM 132
          |  |: :|   :  ||      ||  :  ||:| ||||    :| :|||||||||||:|:|
Qy    62  FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db   133  MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEVYNERVSDLLGKRK 192
          ||     :     ||  |||  ::|  |::: : || :| ::|:||||||||| |||   
Qy   119  MG-----Y-GKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLL-NP- 170

Db   193  KGVKGGGEEVIYDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
          :|| | :|  :|||||||| |  ||: |||:  |:        : || ||||   ::||: 
Qy   171  -STKG---NL--KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db   252  SSRSHAIIMLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSVEGQQFKEAT 311
          |||||| :  | |  |   |||| | || :: | ::|||||||||  : :| :    | 
Qy   224  SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db   312  HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
          ||| ||:||||||  |||| :|  |: ||::|| || | ||||||||||| | ||||:|||
Qy   280  EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db   372  SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
          : :|||||||||||| :|| |  ||:| |||  | |||::: ::|:|
Qy   339  ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385


RESULT 5
ID   US-08-282-845-2      STANDARD;     PRT;    955 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application US/08282845
XX
CC   Sequence 2, Application US/08282845
CC   Patent No. 5719263
CC   GENERAL INFORMATION:
CC   APPLICANT: Reed, Steven G.
CC   TITLE OF INVENTION: A 230Kd Antigen Present in Leishmania
CC   TITLE OF INVENTION: Species
CC   NUMBER OF SEQUENCES: 3
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSE: Immunex Corporation
CC   STREET: 51 University Street
```

CC   CITY: Seattle
CC   STATE: WA
CC   COUNTRY: USA
CC   ZIP: 98101
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: Apple Macintosh
CC   OPERATING SYSTEM: Apple Macintosh Operating System 7.1
CC   SOFTWARE: Microsoft Word for Macintosh 5.1a
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/282,845
CC   FILING DATE:
CC   CLASSIFICATION: 435
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: 08/006,676
CC   FILING DATE: JANUARY 15, 1993
CC   CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Perkins, Patricia Anne
CC   REGISTRATION NUMBER: 34,693
CC   REFERENCE/DOCKET NUMBER: 5004-A
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (206)587-0430
CC   TELEFAX: (206)233-0644
CC   INFORMATION FOR SEQ ID NO: 2:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 955 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ   SEQUENCE   955 AA; 106167 MW; 3654125 CN;

Query Match   17.6%;   Score 953;   DB 1;   Length 955;
Best Local Similarity 43.5%;   Pred. No. 7.47e-65;
Matches 177;   Conservative 90;   Mismatches 111;   Indels 29;   Gaps 18;

Db   13 VKVSVRVRPLNERENNAPEGTKTVTVAAKQAAAVVTVKVLGGSNNSGAAESMGTARRVAQD 72
              :|| ||||||.|| :|: :|:| |: : ::|||: : | : : :
Qy   7 IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA 61

Db   73 FQFDHVFWFSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM 132
              ::|:. :|:|:||:::|||::||||||||||||::|
Qy   62 FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db   133 MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEVYNERVSDLLGKRK 192
              |::|||||||||||||| :|| : | : |:| |:| |
Qy   119 MG-----Y-GKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLL-NP- 170

Db   193 KGVKGGGEEVVDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
              ||::| ||||||||||||| ||:||:|| : : :|:: :|:|:::
Qy   171 -STKG---NL-KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db   252 SSRSHAIIMLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSQVEGGQQFKEAT 311
              ::|:: ||: |: | : | | :::|:|||| |:|||||:
Qy   224 SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db   312 HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
              |:|||| :|||||:||||||||||:||||||||||||||||:||||||||||: : ||
Qy   280 EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db   372 SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
              .:|:|||||||||||||| || ||:||:::|:|:
Qy   339 ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385

RESULT 6
ID   US-08-713-815A-4   STANDARD;   PRT;   411 AA.
XX
AC   xxxxxx
DT
XX
DE   Sequence 4, Application US/08713815A

XX
CC   Sequence 4, Application US/08713815A
CC   Patent No. 5830659
CC   GENERAL INFORMATION:
CC   APPLICANT: Russell J. Stewart
CC   TITLE OF INVENTION: ACTIVE MICROTUBULE-BASED
CC   TITLE OF INVENTION: SEPARATIONS BY KINESINS
CC   NUMBER OF SEQUENCES: 6
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: Thorpe, No. 5830659th & Western, L.L.P.
CC   STREET: 9035 South 700 East, Suite 200
CC   CITY: Sandy
CC   STATE: Utah
CC   COUNTRY: USA
CC   ZIP: 84070
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
CC   COMPUTER: AST Ascentia 900N
CC   OPERATING SYSTEM: DOS 6.22
CC   SOFTWARE: Word Perfect 6.0
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/713,815A
CC   FILING DATE: 13-SEP-1996
CC   CLASSIFICATION: 435
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Alan J. Howarth
CC   REGISTRATION NUMBER: 36,553
CC   REFERENCE/DOCKET NUMBER: T3214/U-2202
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (801)566-6633
CC   TELEFAX: (801)566-0750
CC   INFORMATION FOR SEQ ID NO: 4:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 411 amino acid residues
CC   TYPE: amino acid
CC   TOPOLOGY: linear
SQ   SEQUENCE   411 AA; 45925 MW; 836180 CN;

Query Match   13.0%;   Score 707;   DB 2;   Length 411;
Best Local Similarity 42.3%;   Pred. No. 5.23e-45;
Matches 138;   Conservative 77;   Mismatches 91;   Indels 20;   Gaps 12;

Db   55 FDKVFKPNASQEKVYNEAAKSIVTDVLAGYNGTIFAYGQTGSSGKTHTMEGVIGDSVKQGI 114
              ||| ||: : : ||||:|:|::::| | || |||||||::| : | ::
Qy   71 FDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMG--YGKE--HGV 127

Db   115 IPIRIVNDIFNHIYAM--EVNLEFHIKVSYYEIYMDKIRDLLDVS-KVNLSVHEDKNRVPY 171
              |||| :|: :| | | ::| :|| :|| ||||:||::| :|::   :::
Qy   128 IPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHPSTGPY 187

Db   172 VKGATERFVSSPEDVFEVIEEGKSNRHIAVTNMNEHSSRSHSVFLINVKQENLENQKKL- 230
              ||:| ||: | |: ||:|:::| | :|:|||| ::|||:|:|||:|:::
Qy   188 VEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDEETKMD 247

Db   231 S--GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD---GN--KTH 282
              :   :::|||||||| : :| |||:|||:| :|:| |||::|||| |: |::
Qy   248 TEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKQKKNQ 307

Db   283 -IPYRDSKLTRILQESLGGNARTTVICCSPASFNESETKSTLDFGRRAKTVKNVVCVNE 341
              |||||||| ||:|:|||||:| |: ::|| | ::||| | : : ::|||::| ||
Qy   308 LVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKNHAVVNE 367

Db   342 ELTAEEWKRRYEKEKEKNARLKGKVE 367
              : |:||:: ::::|: ::| :|:
Qy   368 D-PNARMIR-EL-KEELAQLRSKLQ 389

RESULT 7
ID   US-08-713-815A-3   STANDARD;   PRT;   441 AA.
XX
AC   xxxxxx
XX
DT

```
XX
DE    Sequence 3, Application US/08713815A
XX
CC    Sequence 3, Application US/08713815A
CC    Patent No. 5830659
CC    GENERAL INFORMATION:
CC       APPLICANT: Russell J. Stewart
CC       TITLE OF INVENTION: ACTIVE MICROTUBULE-BASED
CC       TITLE OF INVENTION: SEPARATIONS BY KINESINS
CC       NUMBER OF SEQUENCES: 6
CC       CORRESPONDENCE ADDRESS:
CC          ADDRESSEE: Thorpe, No. 5830659th & Western, L.L.P.
CC          STREET: 9035 South 700 East, Suite 200
CC          CITY: Sandy
CC          STATE: Utah
CC          COUNTRY: USA
CC          ZIP: 84070
CC       COMPUTER READABLE FORM:
CC          MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
CC          COMPUTER: AST Ascentia 900N
CC          OPERATING SYSTEM: DOS 6.22
CC          SOFTWARE: Word Perfect 6.0
CC       CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER: US/08/713,815A
CC          FILING DATE: 13-SEP-1996
CC          CLASSIFICATION: 435
CC       ATTORNEY/AGENT INFORMATION:
CC          NAME: Alan J. Howarth
CC          REGISTRATION NUMBER: 36,553
CC          REFERENCE/DOCKET NUMBER: T3214/U-2202
CC       TELECOMMUNICATION INFORMATION:
CC          TELEPHONE: (801)566-6633
CC          TELEFAX: (801)566-0750
CC    INFORMATION FOR SEQ ID NO: 3:
CC       SEQUENCE CHARACTERISTICS:
CC          LENGTH: 441 amino acid residues
CC          TYPE: amino acid
CC          TOPOLOGY: linear
SQ    SEQUENCE  441 AA; 48936 MW; 924366 CN;

Query Match      13.0%;     Score 707;   DB 2;   Length 441;
Best Local Similarity 42.3%;    Pred. No. 5.23e-45;
Matches 138; Conservative  77%; Mismatches  91; Indels 20; Gaps 12;

Db   55 FDKVFKPNASQEKVYNEAAKSIVTDVLAGVNGTIFAYGQTSSGKTHTMEGVIGDSVKQGI 114
         ||  | ||  ||  : :: ||  |||  ||||||::|:  | |    :|:
Qy   71 FDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMWG-YGKE--HGV 127

Db  115 IPRIVNDIFNHIYAM--EVNLEFHIKVSYYEIYMDKIRDLLDVS-KVNLSVHEDKNRVPY 171
         |||| |:|||:| | :    :  || ||   ||:  | | || :  : |
Qy  128 IPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHPSTGPY 187

Db  172 VKGATERFVSSPEDVFEVIEEGKSNRHIAVTNMNEHSSRSHSVFLINVKQENLENQKKL- 230
         | |     || ||  |   : ||:  | || |||| :|||||||| :|     | :
Qy  188 VEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHIDEETKMD 247

Db  231 S---GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD---GN--KTH 282
         :      | || : |||| :| :||  || :|| ||  :  | |:||  :|   :|
Qy  248 TEKVAKISLVDLAGSERAFSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKQKKNQ 307

Db  283 -IPYRDSKLTRILQESLGGNARTTIVICCSPASFNESETKSTLDFGRRAKTVKNVVCVNE 341
          : |||| || :|||| | |     | |  : || || |  || | | |:
Qy  308 LVPYRDSVLFWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKNHAVVNE 367

Db  342 ELTAEEWKRRYEKEKEKNARLKGKVE 367
         :| :|:|   |: :|:||
Qy  368 D--PNARMIR-EL-KEELAQLRSKLQ 389
```

```
RESULT  8
ID    PCT-US94-04496-48      STANDARD;      PRT;     1612 AA.
XX
AC    xxxxxx
```

```
XX
DT
XX
DE    Sequence 48, Application PC/TUS9404496
XX
CC    Sequence 48, Application PC/TUS9404496
CC    GENERAL INFORMATION:
CC       APPLICANT: Croce, Carlo
CC       APPLICANT: Canaani, Eli
CC       TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
CC       TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
CC       TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the ALL-1
CC       NUMBER OF SEQUENCES: 86
CC       CORRESPONDENCE ADDRESS:
CC          ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewitz &
CC          ADDRESSEE: Norris
CC          STREET: One Liberty Place, 46th floor
CC          CITY: Philadelphia
CC          STATE: Pennsylvania
CC          COUNTRY: USA
CC          ZIP: 19103
CC       COMPUTER READABLE FORM:
CC          MEDIUM TYPE: Floppy disk
CC          COMPUTER: IBM PC compatible
CC          OPERATING SYSTEM: PC-DOS/MS-DOS
CC          SOFTWARE: PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER: PCT/US94/04496
CC          FILING DATE:
CC       CLASSIFICATION:
CC       ATTORNEY/AGENT INFORMATION:
CC          NAME: DeLuca Esq., Mark
CC          REGISTRATION NUMBER: 33,229
CC          REFERENCE/DOCKET NUMBER: TJU-1242
CC       TELECOMMUNICATION INFORMATION:
CC          TELEPHONE: (215) 568-3100
CC          TELEFAX: (215) 568-3439
CC    INFORMATION FOR SEQ ID NO: 48:
CC       SEQUENCE CHARACTERISTICS:
CC          LENGTH: 1612 amino acids
CC          TYPE: amino acid
CC          STRANDEDNESS: single
CC          TOPOLOGY: linear
CC       MOLECULE TYPE: protein
CC       HYPOTHETICAL: NO
CC       ANTI-SENSE: NO
SQ    SEQUENCE  1612 AA; 181981 MW; 13031792 CN;

Query Match             3.6%;     Score 196;   DB 4;   Length 1612;
Best Local Similarity 34.5%;    Pred. No. 2.01e-05;
Matches 38; Conservative  24; Mismatches 44; Indels 4; Gaps 4;

Db  383 LPYLVELSPDGSDSRDKPKLYRLQLSVTEVGTEKLDDNS-IQLFGPGIQPHHCDLTNMDG 441
         :|:||||| |   |  :|:|| |  |:| | |: |::|  ||| |:| |||||:|:| :|
Qy  493 MPHLVNLSDDPLLA-E-CLVYNIKPGQTRVGNVNQDTQAEIRLNGSKILKEHCTFENVDN 550

Db  442 VVTVTPRSMDAETYVEGQRISETTMLQSGMKVQFGASHVFKFVDPSQDHA 491
         :|:||||:|   |||   :||  :|:|||| :|: :| ||:|::  ::|
Qy  551 VVTIVPNEKAAVM-VNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEARA 599
```

```
RESULT  9
ID    US-08-545-860D-48      STANDARD;      PRT;     1612 AA.
XX
AC    xxxxxx
XX
DT
XX
DE    Sequence 48, Application US/08545860D
XX
CC    Sequence 48, Application US/08545860D
CC    Patent No. 6040140
CC    GENERAL INFORMATION:
```

```
CC   APPLICANT: Croce, Carlo
CC   APPLICANT: Canaani, Eli
CC   TITLE OF INVENTION:  Diagnostics, Therapeutics and Methods
CC   TITLE OF INVENTION:  for Detection and Treatment of Acute Leukemias
CC   TITLE OF INVENTION:  Resulting from Chromosome Abnormalities in the All-1 Regio
CC   NUMBER OF SEQUENCES:  94
CC   CORRESPONDENCE ADDRESS:
CC     ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewitz &
CC     STREET:  No. 604014Oris
CC     STREET:  One Liberty Place, 46th floor
CC     CITY:  Philadelphia
CC     STATE:  Pennsylvania
CC     COUNTRY:  USA
CC     ZIP:  19103
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE:  Floppy disk
CC     COMPUTER:  IBM PC compatible
CC     OPERATING SYSTEM:  PC-DOS/MS-DOS
CC     SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER:  US/08/545,860D
CC     FILING DATE:  07-MAR-1996
CC     CLASSIFICATION:  435
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  PCT/US94/04496
CC     FILING DATE:  22-APR-1994
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  PCT/US92/10930
CC     FILING DATE:  09-DEC-1992
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  US 08/327,392
CC     FILING DATE:  19-OCT-1994
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  US 08/320,559
CC     FILING DATE:  11-OCT-1994
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  US 08/062,443
CC     FILING DATE:  14-MAY-1993
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  US 07/971,094
CC     FILING DATE:  30-OCT-1992
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  US 07/888,839
CC     FILING DATE:  27-MAY-1992
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  US 07/805,093
CC     FILING DATE:  11-DEC-1991
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME:  DeLuca Esq., Mark
CC     REGISTRATION NUMBER:  33,229
CC     REFERENCE/DOCKET NUMBER:  TJU-1262
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE:  (215) 568-3100
CC     TELEFAX:  (215) 568-3439
CC   INFORMATION FOR SEQ ID NO: 48:
CC   SEQUENCE CHARACTERISTICS:
CC     LENGTH:  1612 amino acids
CC     TYPE:  amino acid
CC     STRANDEDNESS:  single
CC     TOPOLOGY:  linear
CC   MOLECULE TYPE:  protein
CC   HYPOTHETICAL:  NO
CC   ANTI-SENSE:  NO
SQ   SEQUENCE  1612 AA;  181981 MW;  13031792 CN;

Query Match       3.6%;  Score 196;  DB 3;  Length 1612;
Best Local Similarity  34.5%;  Pred. No. 2.0le-05;
Matches  38;  Conservative  24;  Mismatches  44;  Indels  4;  Gaps  4;

Db     383  LPYLVELSPDGSDSRDKPKLYRLQLSVTEVGTEKLDDNS-IQLFGPGIQPHHCDLTNMDG 441
                :  ||::||  :  : :|:::.. .:.|  | :| ::|  :|:  |:|
Qy     493  MPHLVNLSDDPLLA-E-CLVYNIKPGQTRVGNVNQDTQAEIRLNGSKILKEHCTEENVDN 550
```

```
Db     442  VVTVTPRSMDAETYVEGQRISETTMLQSGMKVQFGASHVFKFVDPSQDHA 491
                |||:|  .||  :  :|  |:| :|  |:|  |:||:| |:||:.:|
Qy     551  VVTIVPNEKAAVM-VNGVRIDKPTRLRSGYRILLGDFHIFRFNHPEEARA 599

RESULT 10
ID   US-08-685-871-2         STANDARD;      PRT;    1354 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application US/08685871
XX
CC   Sequence 2, Application US/08685871
CC   Patent No. 6013499
CC   GENERAL INFORMATION:
CC     APPLICANT:  NARUMIYA, Shuh
CC     APPLICANT:  IWAMATSU, Akihiro
CC     TITLE OF INVENTION:  RHO TARGET PROTEIN KINASE P160
CC     NUMBER OF SEQUENCES:  68
CC   CORRESPONDENCE ADDRESS:
CC     ADDRESSEE:  Foley & Lardner
CC     STREET:  3000 K Street, N.W., Suite 500
CC     CITY:  Washington
CC     STATE:  D.C.
CC     COUNTRY:  USA
CC     ZIP:  20007-5109
CC   COMPUTER READABLE FORM:
CC     MEDIUM TYPE:  Floppy disk
CC     COMPUTER:  IBM PC compatible
CC     OPERATING SYSTEM:  PC-DOS/MS-DOS
CC     SOFTWARE:  PatentIn Release #1.0, Version #1.30
CC   CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER:  US/08/685,871
CC     FILING DATE:  24-JUL-1996
CC     CLASSIFICATION:  435
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  JP 8-184102
CC     FILING DATE:  25-JUN-1996
CC   PRIOR APPLICATION DATA:
CC     APPLICATION NUMBER:  JP 7-262553
CC     FILING DATE:  14-SEP-1995
CC   ATTORNEY/AGENT INFORMATION:
CC     NAME:  BENT, Stephen A.
CC     REGISTRATION NUMBER:  29,768
CC     REFERENCE/DOCKET NUMBER:  16887/845
CC   TELECOMMUNICATION INFORMATION:
CC     TELEPHONE:  (202)672-5300
CC     TELEFAX:  (202)672-5399
CC     TELEX:  904136
CC   INFORMATION FOR SEQ ID NO: 2:
CC   SEQUENCE CHARACTERISTICS:
CC     LENGTH:  1354 amino acids
CC     TYPE:  amino acid
CC     TOPOLOGY:  linear
CC   MOLECULE TYPE:  protein
SQ   SEQUENCE  1354 AA;  158174 MW;  8778645 CN;

Query Match       2.2%;  Score 120;  DB 3;  Length 1354;
Best Local Similarity  28.1%;  Pred. No. 2.95e+00;
Matches  41;  Conservative  66;  Mismatches  32;  Indels  7;  Gaps  6;

Db     849  EQYFSTL-YKTQVKELKEIEEKNRENLKKIQELQNEKETLATQLDLAETKAESEQLARG 907
               |:.|||  :|| .|| |:|::: . ::| :|| | ::|.| :|: |:|| :|
Qy     344  EETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGG 403

Db     908  LLEEQYFELTQESKKAASRNRQEITDKDHTVSRLEEANSMLTKDIEILRRE-NEELTEKM 966
             :  ||:  |  | :  |::| |: :: :  ::| .:|:::| |:|:|: ||
Qy     404  PVEESYPPDTPLEKQIVSIQQPDATVK-KMSKAEIVEQ-LNQS-EKLYRDLNQTWEEKL 459

Db     967  KKAEEEYKLEKEEISNLKAAFEKNI 992
             :|||  |:.:|  ::. .: :|||:  :
```

Qy   460 AKTEEIHK-EREAALEELGISIEKGF 484

RESULT 11
ID US-08-785-310A-6   STANDARD;   PRT;   594 AA.
XX
AC xxxxxx
XX
DT
XX
DE
CC Sequence 6, Application US/08785310A
XX
CC Patent No. 5840532
CC GENERAL INFORMATION:
CC   APPLICANT: McKnight, Steven L.
CC   APPLICANT: Russell, David W.
CC   TITLE OF INVENTION: Neuronal PAS Domain Protein
CC   NUMBER OF SEQUENCES: 8
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
CC   STREET: 268 BUSH STREET, SUITE 3200
CC   CITY: SAN FRANCISCO
CC   STATE: CALIFORNIA
CC   COUNTRY: USA
CC   ZIP: 94104
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: Patentin Release #1.0, Version #1.30
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: US/08/785,310A
CC   FILING DATE: 21-JAN-1997
CC   CLASSIFICATION: 536
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: OSMAN, RICHARD A
CC   REGISTRATION NUMBER: 36,627
CC   REFERENCE/DOCKET NUMBER: UTSD:1226
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: (415) 343-4341
CC   TELEFAX: (415) 343-4342
CC INFORMATION FOR SEQ ID NO: 6:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 594 amino acids
CC   TYPE: amino acid
CC   STRANDEDNESS: single
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: peptide
SQ SEQUENCE 594 AA; 63736 MW; 1871071 CN;

Query Match   1.9%; Score 101; DB 2; Length 594;
Best Local Similarity 20.0%; Pred. No. 4.45e+01;
Matches 19; Conservative 27; Mismatches 47; Indels 2; Gaps 2;

Db 174 QVELTGSSVFDYIHPGD-HSEVLEQLGLRAASIGPPTPPSVSSSSSSSSSSSSSLVDTPEIEA 232
         :: | :| : || : :: || || || :: || :: ::
Qy 580 RIILGDFHIFRFNHPEEARAEREQEQSLLRHSVTNSQLGSPAPGRHDRTLSKAGSDADGDS 639

Db 233 SPTEASPAFRAQERS-FFVRMKSTLTKRGLNVKAS 266
         :: | :: | : | | :| : | : |
Qy 640 RSDSPLPHFRGKDSDWFYARREAASAILGLDQKIS 674

RESULT 12
ID PCT-US94-14030A-30   STANDARD;   PRT;   240 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 30, Application PC/TUS9414030A
XX

CC Sequence 30, Application PC/TUS9414030A
CC GENERAL INFORMATION:
CC   APPLICANT: GENETICS INSTITUTE, INC.
CC   APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
CC   TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
CC   NUMBER OF SEQUENCES: 35
CC   CORRESPONDENCE ADDRESS:
CC   ADDRESSEE: GENETICS INSTITUTE, INC.
CC   STREET: 87 CambridgePark Drive
CC   CITY: Cambridge
CC   STATE: Massachusetts
CC   COUNTRY: USA
CC   ZIP: 02140
CC   COMPUTER READABLE FORM:
CC   MEDIUM TYPE: Floppy disk
CC   COMPUTER: IBM PC compatible
CC   OPERATING SYSTEM: PC-DOS/MS-DOS
CC   SOFTWARE: Patentin Release #1.0, Version #1.25
CC   CURRENT APPLICATION DATA:
CC   APPLICATION NUMBER: PCT/US94/14030A
CC   FILING DATE: Herewith
CC   CLASSIFICATION:
CC   PRIOR APPLICATION DATA:
CC   APPLICATION NUMBER: US 08/164,103
CC   FILING DATE: 07-DEC-1993
CC   APPLICATION NUMBER: US 08/217,780
CC   FILING DATE: 25-MAR-1994
CC   APPLICATION NUMBER: US 08/333,576
CC   FILING DATE: 02-NOV-1994
CC   ATTORNEY/AGENT INFORMATION:
CC   NAME: Lazar, Steven R.
CC   REGISTRATION NUMBER: 32,618
CC   REFERENCE/DOCKET NUMBER: 5202D-PCT
CC   TELECOMMUNICATION INFORMATION:
CC   TELEPHONE: 617 498-8260
CC   TELEFAX: 617 876-5851
CC INFORMATION FOR SEQ ID NO: 30:
CC   SEQUENCE CHARACTERISTICS:
CC   LENGTH: 240 amino acids
CC   TYPE: amino acid
CC   TOPOLOGY: linear
CC   MOLECULE TYPE: protein
SQ SEQUENCE 240 AA; 24846 MW; 262270 CN;

Query Match   1.8%; Score 95; DB 4; Length 240;
Best Local Similarity 51.7%; Pred. No. 1.01e+02;
Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Db 91 RRQRTALAGTRGXXGSGGGGGGGGGGGGGGG 119
        | : | |: :|||||||| :|||||| :||
Qy 375 RELKEELAQLRSKLQSSGGGGGGAGGSGG 403

RESULT 13
ID US-08-362-670B-30   STANDARD;   PRT;   240 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 30, Application US/08362670B
XX
CC Sequence 30, Application US/08362670B
CC Patent No. 5658882
CC GENERAL INFORMATION:
CC   APPLICANT: Celeste, Anthony J.
CC   APPLICANT: Wozney, John
CC   APPLICANT: Rosen, Vicki A.
CC   APPLICANT: Wolfman, Neil
CC   APPLICANT: Thomsen, Gerald H.
CC   APPLICANT: Melton, Douglas A.
CC   TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
CC   NUMBER OF SEQUENCES: 37

```
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE: GENETICS INSTITUTE, INC.
CC        STREET: 87 CambridgePark Drive
CC        CITY: Cambridge
CC        STATE: Massachusetts
CC        COUNTRY: USA
CC        ZIP: 02140
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE: Floppy disk
CC        COMPUTER: IBM PC compatible
CC        OPERATING SYSTEM: PC-DOS/MS-DOS
CC        SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: US/08/362,670B
CC        FILING DATE: December 22, 1994
CC        CLASSIFICATION: 514
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME: Lazar, Steven R.
CC        REGISTRATION NUMBER: 32,618
CC        REFERENCE/DOCKET NUMBER: 5202-D
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE: 617 498-8260
CC        TELEFAX: 617 876-5851
CC      INFORMATION FOR SEQ ID NO: 30:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 240 amino acids
CC        TYPE: amino acid
CC        TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ      SEQUENCE 240 AA; 24846 MW; 262270 CN;

Query Match     1.8%;  Score 95;  DB 1;  Length 240;
Best Local Similarity   51.7%;  Pred. No. 1.01e+02;
Matches   15;  Conservative   5;  Mismatches   9;  Indels   0;  Gaps   0;

Db      91 RRQRTALAGTRGXXGSGGGGGGGGGGGGGGG 119
           | : || : | :|||||||:||:|||
Qy     375 RELKEELAQLRSKLQSSGGGGGAGGSGG 403

RESULT  14
ID   US-08-333-576C-30     STANDARD;     PRT;     240 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 30, Application US/08333576C
XX
CC      Sequence 30, Application US/08333576C
CC      Patent No. 6027919
CC      GENERAL INFORMATION:
CC        APPLICANT: Celeste, Anthony J.
CC        APPLICANT: Wozney, John
CC        APPLICANT: Rosen, Vicki A.
CC        APPLICANT: Wolfman, Neil
CC        APPLICANT: Thomsen, Gerald H.
CC        APPLICANT: Melton, Douglas A.
CC        TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
CC        NUMBER OF SEQUENCES: 37
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE: GENETICS INSTITUTE, INC.
CC        STREET: 87 CambridgePark Drive
CC        CITY: Cambridge
CC        STATE: Massachusetts
CC        COUNTRY: USA
CC        ZIP: 02140
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE: Floppy disk
CC        COMPUTER: IBM PC compatible
CC        OPERATING SYSTEM: PC-DOS/MS-DOS
CC        SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
```

```
CC        APPLICATION NUMBER: US/08/333,576C
CC        FILING DATE: No. 6027919ember 2, 1994
CC        CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC        NAME: Lazar, Steven R.
CC        REGISTRATION NUMBER: 32,618
CC        REFERENCE/DOCKET NUMBER: 5202-B
CC      TELECOMMUNICATION INFORMATION:
CC        TELEPHONE: 617 498-8260
CC        TELEFAX: 617 876-5851
CC      INFORMATION FOR SEQ ID NO: 30:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 240 amino acids
CC        TYPE: amino acid
CC        TOPOLOGY: linear
CC      MOLECULE TYPE: protein
SQ      SEQUENCE 240 AA; 24846 MW; 262270 CN;

Query Match     1.8%;  Score 95;  DB 3;  Length 240;
Best Local Similarity   51.7%;  Pred. No. 1.01e+02;
Matches   15;  Conservative   5;  Mismatches   9;  Indels   0;  Gaps   0;

Db      91 RRQRTALAGTRGXXGSGGGGGGGGGGGGGGG 119
           | : || : | :|||||||:||:|||:||
Qy     375 RELKEELAQLRSKLQSSGGGGGGAGGSGG 403

RESULT  15
ID   US-08-860-174A-12     STANDARD;     PRT;     274 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 12, Application US/08860174A
XX
CC      Sequence 12, Application US/08860174A
CC      Patent No. 5989830
CC      GENERAL INFORMATION:
CC        APPLICANT: DAVIS, Paul James
CC        APPLICANT: VAN DER LOGT, Cornelis Paul Erik
CC        APPLICANT: VERHOEIJEN, Martine Elisa
CC        APPLICANT: WILSON, Steve
CC        TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
CC        TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
CC        NUMBER OF SEQUENCES: 31
CC      CORRESPONDENCE ADDRESS:
CC        ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
CC        STREET: 1100 New York Avenue, N.W.
CC        STREET: 9th Floor, East Tower
CC        CITY: WASHINGTON, D.C.
CC        STATE:
CC        COUNTRY: UNITED STATES
CC        ZIP: 20005-3918
CC      COMPUTER READABLE FORM:
CC        MEDIUM TYPE: Floppy disk
CC        COMPUTER: IBM PC compatible
CC        OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text
CC        SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
CC      CURRENT APPLICATION DATA:
CC        APPLICATION NUMBER: US/08/860,174A
CC        FILING DATE: June 16, 1997
CC      PRIOR APPLICATION DATA:
CC        APPLICATION NUMBER: EP 95307332.7
CC        FILING DATE: October 16, 1995
CC        APPLICATION NUMBER: PCT/EP/96/03605
CC        FILING DATE: August 14, 1996
CC      INFORMATION FOR SEQ ID NO: 12:
CC      SEQUENCE CHARACTERISTICS:
CC        LENGTH: 274 amino acids
CC        TYPE: amino acid
CC        TOPOLOGY: linear
CC      MOLECULE TYPE: protein
```

Mon Aug 21 15:58:34 2000

US-09-235-416-1.rai

SQ   SEQUENCE   274 AA; 29378 MW; 452028 CN;

Query Match   1.8%;   Score 98; DB 2; Length 274;
Best Local Similarity   41.3%;   Pred. No. 6.74e+01;
Matches   19; Conservative   10; Mismatches 16; Indels   1; Gaps   1;

Db   141 SSGGGGSGGGGSGGGGSAG-SAQVQLQQSGPELVKPGASVKMSCKA 185
         |||||||:|:||||| : ::: |: : |:||: | |:||: ||
Qy   390 SSGGGGGAGGSGGGPVEESYPPDTPLEKQIVSIQQPDATVKKMSKA 435

Search completed: Mon Aug 21 15:42:50 2000
Job time : 19 secs.

THIS PAGE BLANK (USPTO)

```
 _ _ _ _____  _____  _____   _____  _   _
|  \/  |  __ \/ ____||  __ \ / ____|| | | |
| \  / | |__) | (___ | |__) | |     | |_| |
| |\/| |  ___/ \___ \|  _  /| |     |  _  |
| |  | | |     ____) | | \ \| |____ | | | |   (TM)
|_|  |_|_|    |_____/|_|  \_\\_____||_| |_|
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 21 15:38:36 2000; MasPar time 22.31 Seconds
                                        832.478 Million cell updates/sec

Tabular output not generated.

Title: >US-09-235-416-1
Description: (1-784) from US09235416A.pep
Perfect Score: 5422
Sequence: 1 MSGGGNIKVVRVRPFNARE.........ELRQQQAQMEEALKTAKQEF 784

Scoring table: PAM 150
               Gap 11

Searched: 188963 seqs, 2368106 residues

Post-processing: Minimum Match 08
                 Listing first 45 summaries

Database: a-geneseq36
              1:geneseqp

Statistics: Mean 37.958; Variance 188.559; scale 0.201

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                        SUMMARIES

               %
Result   Query
NO.   Score   Match Length DB   ID   Description   Pred. No.
---- ----- ----- ------ --   -------   -----------   ---------
  1   953   17.6   955   1   W03691   Leishmania chagasi K39   5.54e-67
  2   953   17.6   955   1   R57365   K39 polypeptide of Lei   5.54e-67
  3   791   14.6   2954   1   Y01632   Amino acid sequence of   2.36e-53
  4   707   13.0   411   1   W72745   Drosophila kinesin N-t   2.48e-46
  5   707   13.0   441   1   W72744   Drosophila kinesin N-t   2.48e-46
  6   707   13.0   975   1   W72746   Drosophila kinesin.   2.48e-46
  7   509   9.4   324   1   W70235   Leishmania antigen Lcg   5.83e-30
  8   424   7.8   679   1   W88456   Human kinesin-related   4.74e-23
  9   211   3.9   1201   1   W90345   Drosophila sp. Cos2 pr   2.10e-06
 10   196   3.6   1612   1   W24094   Ras-binding protein p1   2.67e-05
 11   196   3.6   1612   1   R66457   Chimeric ALL-1/AF-6 pr   2.67e-05
 12   183   3.4   1829   1   Y07242   Actin-filament binding   2.35e-04
 13   120   2.2   1354   1   Y07082   Renal cancer associate   5.12e+00
 14   120   2.2   1354   1   W71020   A modified Rho target   5.12e+00
 15   120   2.2   1354   1   W23654   Physiologically active   5.12e+00
 16   113   2.1   350   1   Y13384   Amino acid sequence of   1.44e+01
 17   113   2.1   350   1   W62595   Homo sapiens cerebellu   1.44e+01
 18   106   2.0   235   1   P94260   41kD protein of T. col   3.94e+01
 19   108   2.0   320   1   P94366   41kD protein of T. col   2.96e+01
 20   107   2.0   349   1   W73021   Mouse cysteine-rich se   3.41e+01
 21   108   2.0   350   1   W73016   Human cysteine-rich se   2.96e+01
 22   106   2.0   857   1   W97702   Staphylococcus aureus   3.94e+01
 23   109   2.0   1077   1   R91597   C3G protein.   2.56e+01
```

```
24  106  2.0  1464  1  W42632  Human transcriptional       3.94e+01
25  101  1.9   66   1  W22020  Mutant GCN4 oligomeris      7.96e+01
26  102  1.9  243   1  W60769  Single chain antibody       6.92e+01
27  101  1.9  339   1  R77317  Protein activated lipa      7.96e+01
28  104  1.9  433   1  W18579  Interferon alpha-immun      5.23e+01
29  102  1.9  535   1  W28491  Human p53 protein vari      6.92e+01
30  102  1.9  535   1  W28492  Human p53 protein vari      6.92e+01
31  101  1.9  567   1  R99259  Full-length CD40 bindi      7.96e+01
32  101  1.9  594   1  W68092  Mouse neuronal PAS dom      7.96e+01
33  102  1.9  605   1  W48379  Mus musculus don-1 pol      6.92e+01
34  103  1.9  647   1  W48383  Homo sapiens don-1 pol      6.02e+01
35  104  1.9  1463  1  W99482  Murine NCoA-2 protein.      5.23e+01
36  102  1.9  3054  1  R40841  Translation of TEV lar      6.92e+01
37  100  1.8   50   1  W13633  Apolipoprotein (a) sec      9.15e+01
38  100  1.8  371   1  R83015  Human thyroid transcri      9.15e+01
39  100  1.8  425   1  W80621  S. pneumoniae protein       9.15e+01
40  100  1.8  430   1  Y04287  Streptococcus pneumoni      9.15e+01
41  100  1.8  442   1  R79033  Nuclear inclusion prot      9.15e+01
42  100  1.8  924   1  W60161  Human intercellular ad      9.15e+01
43  100  1.8  924   1  W00931  Human ICAM-4.               9.15e+01
44  100  1.8  924   1  W59006  Human ICAM-4 protein.       9.15e+01
45  100  1.8  924   1  W73511  Human neuronal ICAM-4       9.15e+01
```

                        ALIGNMENTS

```
RESULT  1
ID   W03691 standard; Protein; 955 AA.
AC   W03691;
DT   09-MAR-1997  (first entry)
DE   Leishmania chagasi K39 antigen.
KW   Leishmania chagasi; acidic ribosomal antigen; LcP0;
KW   epitope; K39.
OS   Leishmania chagasi.
PN   WO9633414-A2.
PD   24-OCT-1996.
PF   19-APR-1996; U05472.
PR   21-APR-1995; US-428414.
PA   (CORI-) CORIXA CORP.
PI   Reed SG;
DR   WPI; 96-485884/48.
DR   N-PSDB; T42166.
PT   New Leishmania acidic ribosomal P-protein family poly:peptide - used
PT   to develop prods. for diagnosis, detection and protection against
PT   Leishmania infections.
PS   Disclosure; Page 36-43; 76pp; English.
CC   Compounds including polypeptides that contain at least an epitope of
CC   the L. chagasi acidic ribosomal antigen LcP0 are useful in a variety
CC   of immunoassays for detecting Leishmania infection.  Portions of
CC   LcP0 (T42164) contg. at least the 17 C-terminal amino acids (T42165)
CC   have been found to generate a signal in an ELISA that is equivalent
CC   to that generated by the full length LcP0.  A combination
CC   polypeptide may also be used, comprising an LcP0 epitope along with
CC   an epitope derived from the Leishmania K39 antigen (T42166), pref.
CC   the K39 repeat unit antigen having the sequence given in W03690.
SQ   Sequence 955 AA;

Query Match         17.6%;  Score 953; DB 1; Length 955;
Best Local Similarity 43.5%;  Pred. NO. 5.54e-67;
Matches  177; Conservative  90; Mismatches 111; Indels  29; Gaps  18;

Db   13  VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVVTVKVLGGSNNSGAAESMGTARRVAQD 72
             :||| ||||| | ||  || | | | |: || : : || ::  :    ::| : :||||:| |
Qy    7  IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA 61

Db   73  FQFDHFVEWSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM 132
             || ||||::||::| :: | | :|::| |:|| |:||::|||| ::::||||||||||||| ||
Qy   62  FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db  133  MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEYNERVSDLLGKRK 192
             || |   |:|| ||||| :|||| ::| :: | : | ||::|:|||||| |||| ||||
Qy  119  MG------Y-GKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLL-NP- 170
```

Db 193 KGVKGGGEEVYVDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
         || :: ||||||| :: : : || :|| || :|: :::| |:| |||:|:
Qy 171 -STKG---NL--KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db 252 SSRSHAIIMLLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSVEQQFKEAT 311
         |||| || ||||| :: :||||||||||||| : : ||||||||||||| :: ::||
Qy 224 SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db 312 HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
         |||:||||| |||| |||| :||| || |||||||||||||||||||||||| :|
Qy 280 EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db 372 SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
         : ::|||||||||||| || :| ||: || :||||::| :: :|
Qy 339 ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385

RESULT 2
ID R57365 standard; Protein; 955 AA.
AC R57365;
DT 08-MAR-1995 (first entry)
DE K39 polypeptide of Leishmania chagasi.
KW Parasite; leishmaniasis; infection; repeat unit; vaccine; antigen;
KW Leishmania chagasi.
OS Leishmania chagasi.
PN WO9416331-A.
PD 21-JUL-1994.
PF 10-JAN-1994; U00324.
PR 15-JAN-1993; US-006676.
PA (IASY-) IASYS CORP.
PI Reed SG;
WPI 94-249402/30.
DR N-PSDB: Q70152.
PT Diagnosis of Leishmaniasis - by determining the presence of
PT antibodies that bind to a K39 repeat unit antigen
PS Disclosure; Page 12-15; 28pp; English.
CC The K39 polypeptide comprises a number of repeated units (described
CC in R57366). Detection of antibodies directed against this repeated
CC unit in a patients sample is indicative of leishmaniasis. The
CC antigenic repeat unit can itself be used as a vaccine to protect
CC against infection by a leishmania parasite.
SQ Sequence 955 AA;

Query Match   17.6%;   Score 953;   DB 1; Length 955;
Best Local Similarity 43.5%;   Pred. No. 5.54e-67;
Matches 177;   Conservative 90;   Mismatches 111;   Indels 29;   Gaps 18;

Db   13 VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVTVKVLGGSNNSGAAESMGTARRVAQD 72
         || :|| :| : :| ::: | :|| |:| : : :|:||| || :| |
Qy   7 IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA 61

Db   73 FQFDHVFWSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTVTM 132
         | | ::|: : | | : ||| | : | |: : ||| :||||||||||||| :|
Qy   62 FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db 133 MGADVSALSGEGNGVTPRICLEIFARKASVEAQGHSRWIVELGYVEVYNERVSDLLGKRK 192
         || |::: | | ||| | ::| | :| | ::|| :|||||| |
Qy 119 MG-----Y-GKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEITNERVRDLL-NP- 170

Db 193 KGVKGGGEEVYVDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
         || :: ||||||| :: : : || :|| || :|: :::| |:| |||:|:
Qy 171 -STKG---NL--KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db 252 SSRSHAIIMLLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSVEQQFKEAT 311
         |||| || ||||| :: :||||||||||||| : : ||||||||||||| :: ::||
Qy 224 SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db 312 HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
         |||:||||| |||| |||| :||| || |||||||||||||||||||||||| :|
Qy 280 EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db 372 SALNYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
         : ::|||||||||||| || :| ||: || :||||::| :: :|
Qy 339 ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385

RESULT 3
ID Y01632 standard; Protein; 2954 AA.
AC Y01632;
DT 22-JUN-1999 (first entry)
DE Amino acid sequence of centromere-associated protein-E (CENP-E).
KW CENP-E; centromere-associated protein-E; ATPase activity;
KW plus end-directed microtubule motor activity; chromosome congression;
KW microtubule binding activity; chromosome movement; mitosis;
KW cell proliferation; tumor; metastasis; vascular malfunction;
KW inflammatory disease; immune disease; angiogenesis; hypertension;
KW restenosis; fungal infection; selective herbicide; fungicide;
KW insecticide; plant growth regulator; activator; cancer cell marker.
OS Xenopus sp.
PN WO9913061-A1.
PD 18-MAR-1999.
PF 10-SEP-1998; U19231.
PR 11-SEP-1997; US-058645.
PA (REGC) UNIV CALIFORNIA.
PI Cleveland DW, Goldstein LSB, Sakowicz R, Wood KW;
WPI 99-229233/19.
DR N-PSDB; X26819.
PT Centromere-associated protein-E and related nucleic acid
PS Claim 5; Page 66-67; 77pp; English.
CC The present sequence represents CENP-E (centromere-associated protein-E)
CC of Xenopus. The protein has at least one of plus end-directed microtubule
CC motor activity, ATPase (adenosine triphosphatase) activity and
CC microtubule binding activity. CENP-E is the motor that powers chromosome
CC movement toward microtubule plus ends and is essential for congression
CC of chromosomes during mitosis. Modulators of CENP-E can thus control
CC cell proliferation. Agents that modulate CENP-E activity are lead
CC therapeutic, bioagricultural and diagnostic agents, e.g. for treatment
CC of unwanted cell proliferation (typical of many examples are tumors and
CC metastases; vascular malfunction; inflammatory and immune diseases;
CC angiogenesis; hypertension; restenosis; and fungal infections), also as
CC plant-protection agents (selective herbicides, fungicides and
CC insecticides) and plant growth regulators or activators for improving
CC yields. CENP-E is also a diagnostic marker for dividing cells, including
CC cancer cells.
SQ Sequence 2954 AA;

Query Match   14.6%;   Score 791;   DB 1; Length 2954;
Best Local Similarity 39.9%;   Pred. No. 2.36e-53;
Matches 132;   Conservative 92;   Mismatches 90;   Indels 17;   Gaps 13;

Db   43 KSF-NFDRVFNSHESTSQIYQEIAVPIIRSALQGYNGTIFAYGQTSSGKTYTMMGTPNSL 101
         |:| |||: :|: :: :|:|:|| :|| ||||| |||||||| : |:|: |:|
Qy   66 RSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGYGKEH 125

Db 102 GIIPQAIQEVKIIQEIP-NREFL--LRVSVMEIYNETVKDLLCDDRRKKPLEIREDFNR 158
         |:| :|:| :| :|| |::| : | ::|| |||| : : :||| :|: : |:
Qy 126 GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEITNERVRDLL-NPSTKGNLKVREHPST 184

Db 159 NVYVADLTEELVMVPEHVIQWIKKGEKNRHYGETKMNDHSSRSHTIFRMIVESRDRNDPT 218
         | : ||| | : :| : |::|::| || :| :: : || ||| |||| |
Qy 185 GPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTL-TQKWHDE- 242

Db 219 NSENCDGAVMWSHLNLVDLAGSERASQTGAEGVRLKEGCNINRSLFPILGQVIKKLSD-- 275
         |: | : : :|:|||||||||||: | | :|||| :| :|| | |: |:|||:
Qy 243 ETKM-DTE-KVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSS 300

Db 276 G-QA-GGFINYRDSKLTRILQNSLGGNAKTVIICTITP--VSFDETLSTLQFASTAKHVR 331
         | || ||:|||||: |: | ||||||: || :| | ::|||||||| : :||:
Qy 301 GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK 360

Db 332 NTPHVNEVLDDEALLKRYKREILDLKKQLEN 362
         | ::| :| : || || |:|:|| :::
Qy 361 NHAVVNED-PNARMIRELKEELAQLRSKLQS 390

RESULT 4

ID   W72745 standard; protein; 411 AA.
AC   W72745;
DT   11-JAN-1999 (first entry)
DE   Drosophila kinesin N-terminal 411 amino acid residues.
KW   Drosophila; kinesin; separation; hybridisation; target site;
KW   complex mixture; motor protein; actively transported; separated;
KW   microtubule.
OS   Drosophila sp.
PN   US5830659-A.
PD   03-NOV-1998.
PF   13-SEP-1996; 713815.
PR   13-SEP-1996; US-713815.
PA   (UTAH) UNIV UTAH RES FOUND.
PI   Stewart RJ;
DR   WPI; 98-609236/51.
PT   Separation of selected molecules, e.g. DNA, from complex mixtures -
PT   uses specific apparatus to allow the selected molecule to bind to
PT   motor proteins, and be actively transported and separated away along
PT   micro:tubules.
PS   Claim 3; Column 25-28; 24pp; English.
CC   A method has been developed of separating a selected molecule from a
CC   mixture of molecules. The method comprises: (a) a separation device
CC   comprising a loading reservoir and a receiving reservoir coupled by a
CC   channel with microtubules immobilised on its surface and aligned parallel
CC   to a longitudinal axis of the channel; (b) loading the loading reservoir
CC   with an aqueous solution of the mixture of molecules; (c) adding a motor-
CC   ligand composition and ATP to the solution, where the motor-ligand
CC   comprises, (i) a processive motor capable of attaching to the immobilised
CC   microtubules, and moving in the presence of ATP as source of chemical
CC   energy, and (ii) a ligand coupled to the motor protein, where the ligand
CC   is capable of binding the selected molecule, so that the ligand binds the
CC   selected molecule and the motor protein attaches to the immobilised
CC   microtubules and transports the bound selected molecules along the
CC   receiving reservoir; and (d) removing the selected molecule from the
CC   separation of specific molecules from complex mixtures. The molecule to
CC   be separated (e.g. DNA) binds to the motor protein due to the presence of
CC   the specific binding ligands. Activation of these enables them to travel
CC   down a preformed channel in a specially made piece of apparatus. They can
CC   then be removed easily without contaminants of other mixture particles.
CC   The present sequence represents the N-terminal 411 amino acid residues of
CC   Drosophila kinesin for use in the method of the invention.
SQ   Sequence 411 AA;

Query Match   13.0%; Score 707; DB 1; Length 411;
Best Local Similarity   42.3%; Pred. No. 2.48e-46;
Matches 138; Conservative   77; Mismatches   91; Indels   20; Gaps   12;

Db   55 FDKVFKPNASQEKVYNEAAKSIVTDVLAGYNGTIFAYGQTSSGKTHTMEGVIGDSVKQGI 114
             ||| || : ::   :.: ||| ||||||:|||:|||| :|:
Qy   71 FDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG-YGKE--HGV 127

Db   115 IPRIVNDIFNHIYAM--EVNLEFHIKVSYYEIYMDKIRDLLDVS-KVNLSVHEDKNRVPY 171
             |||| :| :|: || :| || | || :||||: :||||| | |
Qy   128 IPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHPSTGPY 187

Db   172 VKGATERFVSSPEDVFEVIEEGKSNRHIAVTNMNEHSSRSHSVFLINVKQENLENQKKL- 230
             ||||| || | |||:|| | :: ||
Qy   188 VEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDEETKMD 247

Db   231 S---GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD---GN--KTH 282
             :   :||||||| ||| :|| | :| : :||||:| : :|:||| ||| :|
Qy   248 TEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKQKKNQ 307

Db   283 -IPYRDSKLTRILQESLGGNARTTIVICCSPASFNESETKSTLDFGRRAKTVKNVVCVNE 341
              :||||| |:| :||||| ||| ||| :|| ||
Qy   308 LVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLSTLRYADSAKRIKNHAVVNE 367

Db   342 ELTAEEWKRRYEKEKEKNARLKGKVE 367
             : :: :: | | |:|:::
Qy   368 D--PNARMIR-EL-KEELAQLRSKLQ 389

RESULT   5
ID   W72744 standard; protein; 441 AA.
AC   W72744;
DT   11-JAN-1999 (first entry)
DE   Drosophila kinesin N-terminal 441 amino acid residues.
KW   Drosophila; kinesin; separation; hybridisation; target site;
KW   complex mixture; motor protein; actively transported; separated;
KW   microtubule.
OS   Drosophila sp.
PN   US5830659-A.
PD   03-NOV-1998.
PF   13-SEP-1996; 713815.
PR   13-SEP-1996; US-713815.
PA   (UTAH) UNIV UTAH RES FOUND.
PI   Stewart RJ;
DR   WPI; 98-609236/51.
PT   Separation of selected molecules, e.g. DNA, from complex mixtures -
PT   uses specific apparatus to allow the selected molecule to bind to
PT   motor proteins, and be actively transported and separated away along
PT   micro:tubules.
PS   Claim 3; Column 23-26; 24pp; English.
CC   A method has been developed of separating a selected molecule from a
CC   mixture of molecules. The method comprises: (a) a separation device
CC   comprising a loading reservoir and a receiving reservoir coupled by a
CC   channel with microtubules immobilised on its surface and aligned parallel
CC   to a longitudinal axis of the channel; (b) loading the loading reservoir
CC   with an aqueous solution of the mixture of molecules; (c) adding a motor-
CC   ligand composition and ATP to the solution, where the motor-ligand
CC   comprises, (i) a processive motor capable of attaching to the immobilised
CC   microtubules, and moving in the presence of ATP as source of chemical
CC   energy, and (ii) a ligand coupled to the motor protein, where the ligand
CC   is capable of binding the selected molecule, so that the ligand binds the
CC   selected molecule and the motor protein attaches to the immobilised
CC   microtubules and transports the bound selected molecules along the
CC   receiving reservoir; and (d) removing the selected molecule from the
CC   separation of specific molecules from complex mixtures. The molecule to
CC   be separated (e.g. DNA) binds to the motor protein due to the presence of
CC   the specific binding ligands. Activation of these enables them to travel
CC   down a preformed channel in a specially made piece of apparatus. They can
CC   then be removed easily without contaminants of other mixture particles.
CC   The present sequence represents the N-terminal 441 amino acid residues of
CC   Drosophila kinesin for use in the method of the invention.
SQ   Sequence 441 AA;

Query Match   13.0%; Score 707; DB 1; Length 441;
Best Local Similarity   42.3%; Pred. No. 2.48e-46;
Matches 138; Conservative   77; Mismatches   91; Indels   20; Gaps   12;

Db   55 FDKVFKPNASQEKVYNEAAKSIVTDVLAGYNGTIFAYGQTSSGKTHTMEGVIGDSVKQGI 114
             ||| || : ::   :.: ||| ||||||:|||:|||| :|:
Qy   71 FDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG-YGKE--HGV 127

Db   115 IPRIVNDIFNHIYAM--EVNLEFHIKVSYYEIYMDKIRDLLDVS-KVNLSVHEDKNRVPY 171
             |||| :| :|: || :| || | || :||||: :||||| | |
Qy   128 IPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHPSTGPY 187

Db   172 VKGATERFVSSPEDVFEVIEEGKSNRHIAVTNMNEHSSRSHSVFLINVKQENLENQKKL- 230
             ||||| || | |||:|| | :: ||
Qy   188 VEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDEETKMD 247

Db   231 S---GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD---GN--KTH 282
             :   :||||||| ||| :|| | :| : :||||:| : :|:||| ||| :|
Qy   248 TEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKQKKNQ 307

Db   283 -IPYRDSKLTRILQESLGGNARTTIVICCSPASFNESETKSTLDFGRRAKTVKNVVCVNE 341
              :||||| |:| :||||| ||| ||| :|| ||
Qy   308 LVPYRDSVLTWLLKDSLGGNSMTAMIAISPADINFEETLSTLRYADSAKRIKNHAVVNE 367

Db   342 ELTAEEWKRRYEKEKEKNARLKGKVE 367
             : :: :: | | |:|:::
Qy   368 D--PNARMIR-EL-KEELAQLRSKLQ 389

RESULT 6
ID   W72746 standard; Protein; 975 AA.
AC   W72746;
DT   11-JAN-1999 (first entry)
DE   Drosophila kinesin.
KW   Drosophila; kinesin; separation; hybridisation; target site;
KW   complex mixture; motor protein; actively transported; separated;
KW   microtubule.
OS   Drosophila sp.
FH   Key   Location/Qualifiers
FT   Misc_difference 557
FT   /note="encoded by AAT"
PN   US5830659-A.
PD   03-NOV-1998.
PF   13-SEP-1996; 713815.
PR   13-SEP-1996; US-713815.
PA   (UTAH) UNIV UTAH RES FOUND.
PI   Stewart RJ;
DR   WPI; 98-609236/51.
DR   N-PSDB; V67162.
PT   Separation of selected molecules, e.g. DNA, from complex mixtures -
PT   uses specific apparatus to allow the selected molecule to bind to
PT   motor proteins, and be actively transported and separated away along
PT   micro:tubules.
PS   Disclosure; Column 17-24; 24pp; English.
CC   A method has been developed of separating a selected molecule from a
CC   mixture of molecules. The method comprises: (a) a separation device
CC   comprising a loading reservoir and a receiving reservoir coupled by a
CC   channel with microtubules immobilised on its surface and aligned parallel
CC   to a longitudinal axis of the channel; (b) loading the loading reservoir
CC   with an aqueous solution of the mixture of molecules; (c) adding a motor-
CC   ligand composition and ATP to the solution, where the motor-ligand
CC   comprises, (1) a processive motor capable of attaching to the immobilised
CC   microtubules, and moving in the presence of ATP as source of chemical
CC   energy, and (ii) a ligand coupled to the motor protein, where the ligand
CC   is capable of binding the selected molecule, so that the ligand binds the
CC   selected molecule and the motor protein attaches to the immobilised
CC   microtubules and transports the bound selected molecules along the
CC   receiving reservoir; and (d) removing the selected molecule from the
CC   receiving reservoir. The method and the system are used for the
CC   separation of specific molecules from complex mixtures. The molecule to
CC   be separated (e.g. DNA) binds to the motor protein due to the presence of
CC   the specific binding ligands. Activation of these enables them to travel
CC   down a preformed channel in a specially made piece of apparatus. They can
CC   then be removed easily without contaminants of other mixture particles.
CC   The present sequence represents Drosophila kinesin from the present
CC   invention.
SQ   Sequence 975 AA;

Query Match   13.0%;   Score 707; DB 1; Length 975;
Best Local Similarity 42.3%;   Pred. No. 2.48e-46;
Matches 138; Conservative 77; Mismatches 91; Indels 20; Gaps 12;

Db   55 FDKVFKPNASQEKVYNEAAKSIVTDVLAGYNGTIFAYGQTSSGKTHTMEGVIGDSVKQGI 114
             ||| |:||| ::: || |:: ||| |||||:||||:|:|| ||: :|
Qy   71 FDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG--YGKE--HGV 127

Db   115 IPRIVNDIFNHIYAM--EVNLEFHIKVSYYEIYMDKIRDLLDVS-KVNLSVHEDKNRVPY 171
             ||||: || || | ::| ||| :||: ||:|| :| || | :| :| |: | |
Qy   128 IPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHPSTGPY 187

Db   172 VKGATERFVSSPEDVFEVIEEGKSNRHIAVTNMNEHSSRSHSVFLINVKQENLENQKKL- 230
             | |:|| ||| :::| ||:| ||||||:::|||||||| |::|| | |:
Qy   188 VEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDEETKMD 247

Db   231 S---GKLYLVDLAGSEKVSKTGAEGTVLDEAKNINKSLSALGNVISALAD--GN-KTH 282
             :: |||||| | ||||: :|:||||:||| ||:|| |:| |
Qy   248 TEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKQKKNQ 307

Db   283 -IPYRDSKLTRILQESLGGNARTTIVICCSPASFNESETKSTLDFGRRAKTVKNVVCVNE 341
             ||| |||| :||::||| | :| ||||:::| | || :|| |:| | |
Qy   308 LVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKNHAVVNE 367

Db   342 ELTAEEWKRRYEKEKEKNARLKGKVE 367
             :: :: | || | || ||:|:|:
Qy   368 D--PNARMIR-EL-KEELAQLRSKLQ 389

RESULT 7
ID   W70235 standard; Protein; 324 AA.
AC   W70235;
DT   13-NOV-1998 (first entry)
DE   Leishmania antigen LcgSp10 protein.
KW   Leishmania antigen; immune response; infection detection; therapy;
KW   humoral response induction; cellular response induction; cancer;
KW   interleukin-12 production.
OS   Leishmania chagasi.
FH   Key   Location/Qualifiers
FT   Misc_difference 79
FT   /note= "unspecified amino acid"
FT   Misc_difference 96
FT   /note= "unspecified amino acid"
PN   WO9835045-A2.
PD   13-AUG-1998.
PF   12-FEB-1998; U03002.
PR   27-AUG-1997; US-920609.
PR   12-FEB-1997; US-799841.
PA   (CORI-) CORIXA CORP.
PI   Campos-Neto A, Dillon DC, Reed SG, Skeiky YA, Webb JR;
DR   WPI; 98-447242/38.
DR   N-PSDB; V47580.
PT   New immunogenic fragments of Leishmania antigens and related nucleic
PT   acid, vectors and host cells - are useful for diagnosis, prevention
PT   and treatment of leishmaniasis, also to induce production of
PT   interleukin-12 generally
PS   Claim 13; Page 128-129; 194pp; English.
CC   This sequence encodes a Leishmania antigen (LAg) of the invention,
CC   designated LcgSp10. Compositions and vaccines containing the protein are
CC   used to generate a protective or therapeutic immune response against the
CC   Leishmania species donavani, chagasi, infantum, major, amazonensis,
CC   braziliensis, panamensis, tropica or guayanensis. They can also be used
CC   to detect infection (in a skin test). The compositions induce a humoral
CC   and/or cellular response, specifically of Th1 type, particularly
CC   including induction of interleukin-12 (IL-12) production. They may thus
CC   be used more generally to treat any condition (e.g. bacterial, viral or
CC   protozoal infection, or cancer) which responds to IL-12.
SQ   Sequence 324 AA;

Query Match   9.4%;   Score 509; DB 1; Length 324;
Best Local Similarity 42.2%;   Pred. No. 5.83e-30;
Matches 76; Conservative 47; Mismatches 53; Indels 4; Gaps 4;

Db   145 QGIIPRACTDLFDGLRAKRAKDSDFTYRVEVSYYEIYNEKVFDLIRPQRNTDLRIRNSPN 204
             :|| ||:: : | :: | ||:|| ||||||||||:|: |:|: ||: :|::|:|:
Qy   125 HGVIPRICQDMFRRINELQ-KDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHPS 183

Db   205 SGPFIEGLTWKMVSKE-EDVARVIRKGMQERHTAATKFNDRSSRSHAILFNIVQLSMDD 263
             :|: |: :| ::|: | || :: | :|| |::: : |:|||||| | |||: |
Qy   184 TGPYVEDLA-KLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDE 242

Db   264 SDN-AFQMRSKINLVDLAGSERTGAAGAEGNEFHDGVKINHSLTVLGRVIDRLADLSQNK 322
             || | |:|| :|||||||||| |:| || | ||| |||| |||||| || :|:|
Qy   243 ETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGK 302

RESULT 8
ID   W88456 standard; Protein; 679 AA.
AC   W88456;
DT   10-MAY-1999 (first entry)
DE   Human kinesin-related protein KINRELP.
KW   Kinesin-related protein; KINRELP; human; cancer; apoptosis;
KW   cell proliferation; inflammation; therapy; diagnosis; ss.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   Binding_site   286. .293

```
FT   Domain          424..435
FT                   /note= "ATP-binding P-loop"
FT   Modified_site   9..11
FT                   /note= "kinesin motor domain signature"
FT   Modified_site   106..109
FT                   /note= "Asn is N-glycosylated"
FT   Modified_site   244..247
FT                   /note= "Asn is N-glycosylated"
FT   Modified_site   455..458
FT                   /note= "Asn is N-glycosylated"
FT   Modified_site   127..130
FT                   /note= "potential cAMP- and cAMP-dependent protein
FT                   kinase phosphorylation site"
FT   Modified_site   208..211
FT                   /note= "potential cAMP- and cAMP-dependent protein
FT                   kinase phosphorylation site"
FT   Modified_site   4..7
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   27..30
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   108..111
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   187..190
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   273..280
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   338..341
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   427..430
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   441..444
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   447..450
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   582..585
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   634..637
FT                   /note= "potential casein kinase II phosphorylation
FT                   site"
FT   Modified_site   233..241
FT                   /note= "potential tyrosine kinase phosphorylation
FT                   site"
FT   Modified_site   329..336
FT                   /note= "potential tyrosine kinase phosphorylation
FT                   site"
FT   Modified_site   58..60
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   73..75
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   88..90
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   126..128
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   246..248
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   266..268
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"

FT   Modified_site   290..292
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   301..303
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   344..346
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   392..394
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   405..407
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   516..518
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   608..610
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
FT   Modified_site   656..658
FT                   /note= "potential protein kinase C phosphorylation
FT                   site"
PN   WO9858064-A1.
PD   23-DEC-1998.
PF   19-JUN-1998; U12856.
PR   19-JUN-1997; US-878865.
PA   (INCY-) INCYTE PHARM INC.
PI   Au-Young J, Lal P, Shah P;
DR   WPI; 99-080955/07.
DR   N-PSDB: X06945.
PT   New kinesin-related protein - useful for treating cancer and
PT   inflammation
PS   Claim 1; Page 49-50; 72pp; English.
CC   This is the amino acid sequence of human kinesin-related protein
CC   (KINRELP), as deduced from a consensus cDNA sequence (see X06945).
CC   KINRELP shares 97% identity with murine kinesin-related protein
CC   KIF2.   The invention provides expression vectors, host cells,
CC   agonists, antibodies and antagonists, as well as methods for
CC   treating disorders associated with expression of KINRELP.  KINRELP
CC   and its agonists are used to stimulate cell proliferation, and to
CC   treat a disorder associated with increased apoptosis.   Antagonists
CC   are used to treat cancer and inflammation (claimed).  Particularly,
CC   KINRELP or its fragment may be used to prevent or treat: AIDS and
CC   other infectious or genetic immunodeficiencies; neurodegenerative
CC   diseases such as Alzheimer's disease, Parkinson's disease,
CC   amyotrophic lateral sclerosis, retinitis pigmentosis and
CC   cerebellar degeneration; myelodysplastic syndromes such as aplastic
CC   anaemia; ischemic diseases such as myocardial infarction, stroke
CC   and reperfusion injury; toxin-induced diseases such as
CC   alcohol-induced liver damage, cirrhosis and lathyrism; wasting
CC   diseases such as cachexia and osteoporosis; viral infections such
CC   as those caused by hepatitis B and C; and diseases associated with
CC   inflammation including adult respiratory distress syndrome,
CC   allergy, asthma, arteriosclerosis, bronchitis, emphysema,
CC   hypereosinophilia, myocardial or pericardial inflammation,
CC   rheumatoid arthritis, Addison's disease, AIDS, anaemia,
CC   atherosclerosis, diseases of the digestive system, atopic
CC   dermatitis, dermatomyositis, diabetes mellitus, glomerulonephritis,
CC   gout, Grave's disease, lupus erythematosus, multiple sclerosis,
CC   myasthenia gravis, osteoarthritis, osteoporosis, pancreatitis,
CC   polycystic kidney disease, polymyositis, scleroderma, Sjorgren's
CC   syndrome, autoimmune thyroiditis, complications of cancer,
CC   extracorporeal circulation, viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections and trauma.
SQ   Sequence  679 AA;

Query Match    7.8%;  Score 424;  DB 1;  Length 679;
Best Local Similarity 35.7%;  Pred. No. 4.74e-23;
Matches  133;  Conservative  80;  Mismatches 121;  Indels  39;  Gaps  23;
```

```
Db   199  VCVRKRPLNKKETQMKDLDVITIPSKDVVMHEP--KQKVDLT-R-YLENQ-TFRFD--Y   251
          | ||| ||  :| :|   : :: :   :: : :|   :|  ::: : :|||  |  |
Qy     9  VVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAFAFDRSY   68

Db   252  -AFDDSAPN---EMVYRFTARPLVETIFERGMATC-FAYGQTGSGKTHTMGGDFSGKNQ   305
          :|| ::|     | :: || | |  :|  |  -  | ||||||||||: |:
Qy    69  WSFDKNAPNYARQEDLFQDLGVPLLDNAF-KGYNNCIFAYGQTGSGKSYSMMG-Y-GKEH  125

Db   306  DCSKGIYALAARDVFLMLKKPNYKKLELQVYATFFEIYSGKVFDLLNRKTK--LRVLEDG  363
          | | |   :: |  :| | : ||   :| :: :|  :| : | ::|| :| |  LRVLEDG
Qy   126  GVIPRICQDMFRRIN-ELQK-D-KNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREHP  182

Db   364  KQQVQVVGLQEREKCVEDVLKLIDIGNSCRTSGQTSANAHSSRSHAVFQIILRRK--G-  420
          : | ::  :  |  : ::| :| | : || | :|     ||||||| : :  :|
Qy   183  STGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDE  242

Db   421  --KLH-----GKFSLIDLAGNERGADTSSADRQTRLEGAEINKSLLALKECIRAL-----G  469
          :|        ||| | |||||  :|      | ::  |  |  :|  | :
Qy   243  ETKMDTEKVAKISLVDLAGSERATSTGATGARLK-EGAEINRSLSTLGRVIAALADMSSG  301

Db   470  R-NKPH-TPFRASKLTQVLRDSFIGENSRTCMIATISPGMASCENTLNTLRYANRVKELT  527
          : :: ||  |||  | || |  | |||| : | ||  ||| :|| |:|||||:: | :|
Qy   302  KQKKNQLVPYRDSVLTWLLKDS-LGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK  360

Db   528  VDPTAAGDVRPIM  540
          | :|    || :|
Qy   361  NHAVVNEDPNARM  373


RESULT  9
ID   W90345  standard; Protein; 1201 AA.
AC   W90345;
DT   25-MAY-1999  (first entry)
DE   Drosophila sp. Cos2 protein.
KW   Cos2; fruitfly; Costal2; kinesin-related; hedgehog; modulate; mapping;
KW   gene therapy, prophylactic; therapeutic; treatment; cancer; meningioma;
KW   basal cell carcinoma; glioma; medulloblastoma; breast carcinoma;
KW   hedgehog signalling pathway; infertility treatment; contraception;
KW   developmental abnormalities; immunogen.
OS   Drosophila sp.
FH   Key             Location/Qualifiers
FT   Region          1..450
FT                   /note= "N-terminal region"
FT   Binding_site    174..181
FT                   /note= "Putative nucleotide binding motif N1 (P-loop)"
FT   Binding_site    237..239
FT                   /note= "Putative microtubule binding motif"
FT   Binding_site    265..269
FT                   /note= "Putative nucleotide binding motif N2"
FT   Binding_site    298..303
FT                   /note= "Putative nucleotide binding motif N3"
FT   Binding_site    340..345
FT                   /note= "Putative microtubule binding motif"
FT   Region          643..990
FT                   /note= "central region"
FT   Region          643..650
FT                   /note= "heptad repeat"
FT   Region          651..656
FT                   /note= "heptad repeat"
FT   Region          657..663
FT                   /note= "heptad repeat"
FT   Region          664..670
FT                   /note= "heptad repeat"
FT   Region          671..677
FT                   /note= "heptad repeat"
FT   Region          678..684
FT                   /note= "heptad repeat"
FT   Region          685..691
FT                   /note= "heptad repeat"
FT   Region          692..698
FT                   /note= "heptad repeat"
FT   Region          706..712
FT                   /note= "heptad repeat"
FT   Region          713..719
FT                   /note= "heptad repeat"
FT   Region          720..726
FT                   /note= "heptad repeat"
FT   Region          727..733
FT                   /note= "heptad repeat"
FT   Region          740..746
FT                   /note= "heptad repeat"
FT   Region          747..753
FT                   /note= "heptad repeat"
FT   Region          754..760
FT                   /note= "heptad repeat"
FT   Region          761..767
FT                   /note= "heptad repeat"
FT   Region          771..777
FT                   /note= "heptad repeat"
FT   Region          778..784
FT                   /note= "heptad repeat"
FT   Region          785..791
FT                   /note= "heptad repeat"
FT   Region          792..798
FT                   /note= "heptad repeat"
FT   Region          799..805
FT                   /note= "heptad repeat"
FT   Region          806..812
FT                   /note= "heptad repeat"
FT   Region          833..839
FT                   /note= "heptad repeat"
FT   Region          840..846
FT                   /note= "heptad repeat"
FT   Region          847..853
FT                   /note= "heptad repeat"
FT   Misc_difference 857
FT                   /label= unknown
FT   Region          870..876
FT                   /note= "heptad repeat"
FT   Region          877..883
FT                   /note= "heptad repeat"
FT   Region          884..890
FT                   /note= "heptad repeat"
FT   Region          905..911
FT                   /note= "heptad repeat"
FT   Region          912..918
FT                   /note= "heptad repeat"
FT   Region          919..925
FT                   /note= "heptad repeat"
FT   Region          926..932
FT                   /note= "heptad repeat"
FT   Region          963..969
FT                   /note= "heptad repeat"
FT   Region          970..976
FT                   /note= "heptad repeat"
FT   Region          977..983
FT                   /note= "heptad repeat"
FT   Region          984..990
FT                   /note= "heptad repeat"
FT   Region          1050..1201
FT                   /note= "C-terminal region"
PN   WO900403-A1.
PD   07-JAN-1999.
PF   18-JUN-1998; U12820.
PF   30-JUN-1997; US-051347.
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PI   Scott MP, Sisson JC;
DR   WPI; 99-095669/08.
DR   N-PSDB; V72291.
PT   New costal2 (cos2) gene - used to identify homologous and related
PT   proteins
PS   Claim 3; Page 31-34; 42pp; English.
CC   This sequence represents a Cos2 gene isolated from Drosophila sp. The
CC   Cos2 protein (also known as the Costal2 protein) is a kinesin-related
CC   protein that accumulates preferentially in cells capable of responding to
CC   hedgehog. This sequence is used to identify homologous or related genes,
```

CC   in producing compositions that modulate the expression or function of the
CC   encoded protein, for gene therapy, mapping functional regions of the
CC   protein and in studying associated physiological pathways. In addition,
CC   modulation of Cos2 gene activity is used for prophylactic and therapeutic
CC   purposes such as treatment of cancers (e.g. basal cell carcinoma, glioma,
CC   medulloblastoma, meningioma, breast carcinoma), hedgehog signalling
CC   pathway function, identification of cell type based on expression,
CC   infertility treatment, contraception, developmental abnormalities. The
CC   encoded protein is useful as an immunogen for producing specific
CC   antibodies, in screening for biologically active agents that modulate
CC   Cos2 function and that act in the hedgehog signalling pathway. The
CC   protein can also be used for therapeutic and prophylactic purposes.
SQ   Sequence 1201 AA;

Query Match   3.9%; Score 211; DB 1; Length 1201;
Best Local Similarity 31.6%; Pred. No. 2.10e-06;
Matches   50; Conservative   42; Mismatches   56; Indels 10; Gaps   7;

Db   268 AHTLFTLTLEQQWVSKEGLLQ-HRLSTASFSDLCGTERGGDQ-PPGRPL-D-A----GLC 319
          :!:.!!!!!!! :! :!! ! !!.:: :: ! !! !!!   ! :!
Qy   227 SHAVFTLTLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLS 286

Db   320 MLEQVISTLTD-PGLMYGVNGNIPYGQTLITTLLKDSFGGRAQTLVILCVSPLEEHLPET 378
          !:!!!:!! .:: . : :!! !!! !!!!!!!!:!.:! .!! :!! :: !!
Qy   287 TLGRVIAALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSWTAMIAAISPADINFEET 346

Db   379 LGNLQFAFKVQCVRNFVIMNTYSDDNTMIVQPAEPVPE 416
          ! :!:! ! :!   ! ! .:! ! !!
Qy   347 LSTLRYADSAKRIKNHAVVNE-DPNARMIRELKEELAQ 383

RESULT 10
ID   W24094 standard; Protein; 1612 AA.
AC   W24094;
DT   20-NOV-1997 (first entry)
DE   Ras-binding protein p180.
KW   Ras; binding protein; p180; modified AF-6; recombinant production;
KW   inhibit; tumour formation; metastasis; screen.
OS   Homo sapiens.
PN   J09191879-A.
PD   29-JUL-1997.
PF   07-DEC-1995; 319281.
PR   15-NOV-1995; JP-297109.
PA   (KIRI) KIRIN BREWERY KK.
DR   WPI; 97-429179/40.
DR   N-PSDB; T85642.
PT   Active type Ras-binding protein p180 - used to inhibit tumour
PT   formation
CC   Claim 1; Page 11-17; 24pp; Japanese.
CC   This sequence is a protein, p180, which comprises a modified AF-6 amino
CC   acid sequence which has active Ras-binding activity. Host cells
CC   transformed with the cDNA can be cultured for the recombinant production
CC   of p180. p180 inhibits tumour formation or metastasis. It can also be
CC   used to screen for substances that inhibit binding between active type
CC   Ras protein and AF-6.
SQ   Sequence 1612 AA;

Query Match   3.6%; Score 196; DB 1; Length 1612;
Best Local Similarity 34.5%; Pred. No. 2.67e-05;
Matches   38; Conservative   24; Mismatches   44; Indels 4; Gaps   4;

Db   383 LPYLVELSPDGSDSRDKPKLYRLQLSVTEVGTEKLDDNS-IQLFGPGIQPHHCDLTNMDG 441
          !:!!!:!! :!   :! :!! !!:::!!!! !:! !! !! ::!:! ::!!.!:!
Qy   493 MPHLVNLSDDPLLA-E-CLVYNIKPGQTRVGNVNQDTQAEIRLNGSKILKEHCTFENVDN 550

Db   442 VVTVTPRSMDAETYVEGQRISETTMLQSGMKVQFGASHVFKFVDPSQDHA 491
          !!!! :! ! :!! :::!!!!!!!! :! :! :!! :!! !!!!! :!!
Qy   551 VVTIVPNEKAAVM-VNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEARA 599

RESULT 11
ID   R66457 standard; Protein; 1612 AA.
AC   R66457;
DT   18-AUG-1995 (first entry)
DE   Chimeric ALL-1/AF-6 protein.
KW   Acute lymphoblastic leukaemia; acute nonlymphoblastic leukaemia;
KW   chromosomal translocation; abnormality; detection;
OS   Homo sapiens.
PN   WO942693O-A.
PD   24-NOV-1994.
PF   22-APR-1994; U04496.
PR   14-MAY-1993; US-062443.
PA   (UVJE-) UNIV JEFFERSON THOMAS.
PI   Canaani E, Croce C;
DR   WPI; 95-006818/01.
PT   New acute lymphocytic leukaemia gene prods. - used for the
PT   diagnosis and treatment of leukaemias, partic. acute
PT   lymphoblastic or nonlymphoblastic leukaemia
PS   Claim 34; Page 121-127; 207pp; English.
CC   Monoclonal antibodies which bind to at least part of the chimeric
CC   ALL-1/AF-6 proteins R66456-R66458 are claimed. The antibodies are
CC   useful for diagnosing acute lymphoblastic and non-lymphoblastic
CC   leukaemia.
SQ   Sequence 1612 AA;

Query Match   3.6%; Score 196; DB 1; Length 1612;
Best Local Similarity 34.5%; Pred. No. 2.67e-05;
Matches   38; Conservative   24; Mismatches   44; Indels 4; Gaps   4;

Db   383 LPYLVELSPDGSDSRDKPKLYRLQLSVTEVGTEKLDDNS-IQLFGPGIQPHHCDLTNMDG 441
          !:!!!:!! :!   :! :!! !!:::!!!! !:! !! !! ::!:! ::!!.!:!
Qy   493 MPHLVNLSDDPLLA-E-CLVYNIKPGQTRVGNVNQDTQAEIRLNGSKILKEHCTFENVDN 550

Db   442 VVTVTPRSMDAETYVEGQRISETTMLQSGMKVQFGASHVFKFVDPSQDHA 491
          !!!! :! ! :!! :::!!!!!!!! :! :! :!! :!! !!!!! :!!
Qy   551 VVTIVPNEKAAVM-VNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEARA 599

RESULT 12
ID   Y07242 standard; protein; 1829 AA.
AC   Y07242;
DT   06-JUL-1999 (first entry)
DE   Actin-filament binding protein 1-Afadin.
KW   Actin-binding protein 1-Afadin; rat; foetal brain tissue; infiltration;
KW   metastasis; carcinoma; diagnosis.
OS   Rattus sp.
PN   EP-905239-A2.
PD   31-MAR-1999.
PF   21-SEP-1998; 307643.
PR   22-SEP-1997; JP-257043.
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA   (OBAI/) OBAISHI H.
PA   (WADA/) WADA M.
DR   WPI; 99-206773/18.
PT   New actin filament-binding protein 1-Afadin is useful for diagnosing
PT   and treating carcinomas
PS   Claim 1; Page 8-18; 27pp; English.
CC   This sequence represent an actin-binding protein 1-Afadin derived from
CC   rat foetal brain tissue. The isolated protein has a molecular weight
CC   of around 205 kD. 1-Afadin and its derivatives may be used to clarify
CC   the mechanisms of infiltration and metastasis of carcinoma useful in
CC   diagnosing or treating carcinomas.
SQ   Sequence 1829 AA;

Query Match   3.4%; Score 183; DB 1; Length 1829;
Best Local Similarity 35.6%; Pred. No. 2.35e-04;
Matches   32; Conservative   19; Mismatches   37; Indels 2; Gaps   2;

Db   418 LYRLQLSVTEVGTEKFDDNS-IQLFGPGIQPHHCDLTNMDGVVTVTPRSMDAETYVDGQR 476
          :!: !:!! :! :!!:! !! !! ::!:! ::!!.!:! !!!! :! ! :!! : :!
Qy   510 VNIKPGQTRVGNVNQDTQAEIRLNGSKILKEHCTFENVDNVVTIVPNEKAAVM-VNGVR 568

Db   477 ISETTMLQSGMRLQFGTSHVFKFVDPIQDH 506
          !!!!!!! :! : !!! :!! !!!!! :!!
Qy   569 IDKPTRLRSGYRIILGDFHIFRFNHPEEAR 598

RESULT 13
ID   Y07082 standard; Protein; 1354 AA.
AC   Y07082;
DT   02-JUL-1999 (first entry)
DE   Renal cancer associated antigen precursor sequence.
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer.
OS   Homo sapiens.
PN   WO9904265-A2.
PD   28-JAN-1999.
PF   15-JUL-1998; U14679.
PR   22-JUN-1998; US-102322.
PR   17-JUL-1997; US-896164.
PR   10-OCT-1997; US-061599.
PR   10-OCT-1997; US-061765.
PR   10-OCT-1997; US-948705.
PR   11-OCT-1997; GB-021697.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Chen Y, Gout I, Gure A, OHare M, Obata Y, Old LJ,
PI   Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E,
PI   Tureci O;
DR   WPI; 99-132448/11.
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
PS   Disclosure; Page 491-494; 787pp; English.
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
SQ   Sequence 1354 AA;

Query Match   2.2%; Score 120; DB 1; Length 1354;
Best Local Similarity   28.1%;   Pred. No. 5.12e+00;
Matches   41; Conservative   32; Mismatches   66; Indels   7; Gaps   6;

Db   849 EQYFSTL-YKTQVKELKEEIEEKNRENLKKIQELQNEKETLATQLDLAETKAESEQLARG 907
         |::||| | |:: |: ::|:.:||:||::|::|||:| :: : . :::   :|
Qy   344 EETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGG 403

Db   908 LLEEQYFELTQESKKAASRNRQEITDKDHTVSRLEEANSMLTKDIEILRRE-NEELTEKM 966
         :||| | | | :: | | |: || : |||:| : |:|: | | |: :||: ||
Qy   404 PVEESYPPDTPLEKQIVSIQQPDATVK--KMSKAEIVEQ-LNQS-EKLYRDLNQTWEEKL 459

Db   967 KKAEEEYKLEKEEISNLKAAFEKNI 992
         |:|| | || ::| ::|||
Qy   460 AKTEEIHK-EREAALEELGISIEKGF 484

RESULT 14
ID   W71020 standard; Protein; 1354 AA.
AC   W71020;
DT   27-OCT-1998 (first entry)
DE   A modified Rho target protein kinase p160 protein.
KW   protein kinase activity; Rho protein; preparation;
KW   therapeutic composition.
OS   Homo sapiens.
PN   J10191985-A.
PD   28-JUL-1998.
PF   17-JAN-1997; 019870.
PR   17-JAN-1997; JP-019870.
PA   (KIRI) KIRIN BREWERY KK.
DR   WPI; 98-460110/40.
DR   N-PSDB; V42941.
PT   New protein exhibiting protein kinase activity - is not capable of
PT   binding to active Rho protein or its derivative, used, e.g.
PT   therapeutically
PS   Claim 1; Fig 1; 60pp; Japanese.
CC   The present sequence encodes a protein which exhibits protein kinase
CC   activity and is not capable of binding to active Rho protein or its
CC   derivative. The materials may be used for the preparation of
CC   therapeutic compositions.
SQ   Sequence 1354 AA;

Query Match   2.2%; Score 120; DB 1; Length 1354;
Best Local Similarity   28.1%;   Pred. No. 5.12e+00;
Matches   41; Conservative   32; Mismatches   66; Indels   7; Gaps   6;

Db   849 EQYFSTL-YKTQVKELKEEIEEKNRENLKKIQELQNEKETLATQLDLAETKAESEQLARG 907
         |::||| | |:: |: ::|:.:||:||::|::|||:| :: : . :::   :|
Qy   344 EETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGG 403

Db   908 LLEEQYFELTQESKKAASRNRQEITDKDHTVSRLEEANSMLTKDIEILRRE-NEELTEKM 966
         :||| | | | :: | | |: || : |||:| : |:|: | | |: :||: ||
Qy   404 PVEESYPPDTPLEKQIVSIQQPDATVK--KMSKAEIVEQ-LNQS-EKLYRDLNQTWEEKL 459

Db   967 KKAEEEYKLEKEEISNLKAAFEKNI 992
         |:|| | || ::| ::|||
Qy   460 AKTEEIHK-EREAALEELGISIEKGF 484

RESULT 15
ID   W23654 standard; Protein; 1354 AA.
AC   W23654;
DT   13-OCT-1997 (first entry)
DE   Physiologically active protein p160.
KW   Rho binding activity; protein kinase; protein binding; host cell;
KW   recombinant production; human.
OS   Homo sapiens.
PN   J09135683-A.
PD   27-MAY-1997.
PF   25-JUN-1996; 184102.
PR   14-SEP-1995; JP-262553.
PA   (KIRI) KIRIN BREWERY KK.
DR   WPI; 97-335990/31.
DR   N-PSDB; T78203.
PT   Physiologically active protein p160 - has rho binding activity and
PT   protein kinase activity
PS   Claim 6; Page 21-24; 53pp; Japanese.
CC   The present sequence represents the physiologically active protein
CC   p160, which has active Rho protein-binding and protein kinase
CC   activities. The nucleotide sequence can be put into a vector which can
CC   then be used to transform a host cell. The host cell can be cultured
CC   for the recombinant production of the p160 protein. The p160 protein
CC   can be used in a method to screen for compounds that inhibit the
CC   protein kinase activity.
SQ   Sequence 1354 AA;

Query Match   2.2%; Score 120; DB 1; Length 1354;
Best Local Similarity   28.1%;   Pred. No. 5.12e+00;
Matches   41; Conservative   32; Mismatches   66; Indels   7; Gaps   6;

Db   849 EQYFSTL-YKTQVKELKEEIEEKNRENLKKIQELQNEKETLATQLDLAETKAESEQLARG 907
         |::||| | |:: |: ::|:.:||:||::|::|||:| :: : . :::   :|
Qy   344 EETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGG 403

Db   908 LLEEQYFELTQESKKAASRNRQEITDKDHTVSRLEEANSMLTKDIEILRRE-NEELTEKM 966
         :||| | | | :: | | |: || : |||:| : |:|: | | |: :||: ||
Qy   404 PVEESYPPDTPLEKQIVSIQQPDATVK--KMSKAEIVEQ-LNQS-EKLYRDLNQTWEEKL 459

Db   967 KKAEEEYKLEKEEISNLKAAFEKNI 992
         |:|| | || ::| ::|||
Qy   460 AKTEEIHK-EREAALEELGISIEKGF 484

US-09-235-416-1.rag

Mon Aug 21 15:58:33 2000

Search completed: Mon Aug 21 15:39:03 2000
Job time : 27 secs.
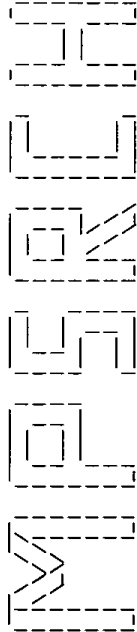
```
M P S R C H (TM)
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh. U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 21 15:39:19 2000; MasPar time 40.22 Seconds
           919.448 Million cell updates/sec

Tabular output not generated.

Title: >US-09-235-416-1
Description: (1-784) from US09235416A.pep
Perfect Score: 5422
Sequence: 1 MSGGGNIKVVVVRPFNARE.........ELRQQAQMEEALKTAKQEF 784

Scoring table: PAM 150
                Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database: pir64
                1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 53.618; Variance 119.927; scale 0.447

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	2217	40.9	1150	1	A55289	kinesin-like protein	0.00e+00
2	2176	40.1	1695	2	A56921	kinesin family protei	0.00e+00
3	1975	36.4	1584	2	JN0114	kinesin-related prote	0.00e+00
4	1971	36.4	1584	1	T15822	kinesin-like protein	0.00e+00
5	1597	29.5	1921	2	T13827	kinesin-73 - fruit fl	1.82e-262
6	1291	23.8	928	2	T16759	hypothetical protein	3.61e-205
7	1008	18.6	699	1	S38982	kinesin-related prote	1.11e-152
8	962	17.7	701	1	B44259	kinesin-related prote	3.24e-144
9	962	17.7	742	1	S58691	kinesin-related prote	3.24e-144
10	953	17.6	955	2	A47334	LcKin kinesin-related	1.46e-142
11	942	17.4	747	2	A57107	kinesin-related prote	1.53e-140
12	900	16.6	294	2	S38983	kinesin-related prote	7.69e-133
13	897	16.5	784	2	A55236	kinesin-related prote	2.73e-132
14	870	16.0	786	2	A53939	kinesin homolog KHP1	2.37e-127
15	857	15.8	1121	2	T13750	kinesin-like protein	5.60e-125
16	850	15.7	1121	2	T13796	kinesin-related prote	1.06e-123
17	846	15.6	1231	2	A54803	microtubule-associate	5.70e-123
18	839	15.5	672	2	S54351	kinesin osm-3 - Caeno	1.08e-121
19	823	15.2	1225	2	A56514	chromokinesin - chick	8.86e-119
20	825	15.2	1226	2	I51617	kinesin-like protein	3.83e-119
21	815	15.0	928	2	T10164	kinesin heavy chain -	2.53e-117
22	799	14.7	2663	2	S28261	centromere protein E	2.06e-114
23	791	14.6	2954	2	T14156	kinesin-related prote	5.84e-113
24	785	14.5	332	2	C48835	kinesin-like protein	7.17e-112
25	762	14.1	805	2	S64238	kinesin-related prote	1.06e-107
26	741	13.7	967	1	A35075	kinesin heavy chain -	6.69e-104
27	731	13.5	1031	1	A38713	kinesin heavy chain -	4.28e-102
28	727	13.4	793	2	JC5831	kinesin-related prote	2.26e-101
29	725	13.4	815	2	A48264	kinesin heavy chain u	5.19e-101
30	725	13.4	843	1	S44868	R05D3.7 protein - Cae	5.19e-101
31	707	13.0	975	1	A31497	kinesin heavy chain -	9.12e-98
32	704	13.0	1032	2	I38510	neuronal kinesin heav	3.16e-97
33	697	12.9	1027	1	S37711	kinesin heavy chain -	5.75e-96
34	687	12.7	147	2	E44259	kinesin-related prote	3.61e-94
35	665	12.3	330	2	B48835	kinesin-like protein	3.21e-90
36	665	12.3	963	1	A41919	kinesin heavy chain -	3.21e-90
37	660	12.2	881	2	I84737	kinesin-like spindle	2.52e-89
38	654	12.1	1056	1	G02157	kinesin heavy chain -	2.99e-88
39	656	12.1	1060	1	A40264	kinesin-related prote	1.31e-88
40	648	12.0	776	2	A53953	kinesin-like protein	3.54e-87
41	653	12.0	1066	1	A48669	kinesin-related prote	4.52e-88
42	652	12.0	1067	2	S33417	kinesin-like protein	6.82e-88
43	649	12.0	1184	2	A34795	kinesin-related prote	2.35e-87
44	627	11.6	754	2	S48020	kinesin-related prote	1.99e-83
45	623	11.5	745	2	S48019	kinesin-related prote	1.03e-82

ALIGNMENTS

RESULT   1

ENTRY   A55289   #type complete
TITLE   kinesin-like protein KIF1B - mouse
ORGANISM   10-Sep-1999 #formal_name Mus musculus #common_name house mouse
DATE   10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
ACCESSIONS   A55289
REFERENCE   A55289
   #authors   Nangaku, M.; Sato-Yoshitake, R.; Okada, Y.; Noda, Y.;
                Takemura, R.; Yamazaki, H.; Hirokawa, N.
   #journal   Cell (1994) 79:1209-1220
   #title   KIF1B, a novel microtubule plus end-directed monomeric motor
                protein for transport of mitochondria.
   #cross-references MUID:95094296
##accession   A55289
##status   preliminary
##molecule_type mRNA
##residues   1-1150 ##label NAN
##cross-references GB:D17577; NID:g407338; PIDN:BAA04503.1;
                   PID:d1005029; PID:g407339
CLASSIFICATION #superfamily kinesin-related protein KIF1B; kinesin motor
                domain homology
KEYWORDS   p-loop
FEATURE
   6-354   #domain kinesin motor domain homology #label KMOT\
   97-104   #region nucleotide-binding motif A (P-loop)
SUMMARY   #length 1150 #molecular-weight 130278 #checksum 5200

Query Match   40.9%;   Score 2217; DB 1; Length 1150;
Best Local Similarity 55.3%;   Pred. No. 0.00e+00;
Matches 340; Conservative 132; Mismatches 114; Indels 29; Gaps 20;

Db   1 MSGAS-VKVAVRVRPFNSRETSKESKCIIQMQGNSTSIINP----K--NP·K---EAPK   48
          |||| :||||||||| :   : : :||: : ::|||| :|| :||
Qy   1 MSGGGNIKVVVVRPFNFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60

Db   49 SFSFDYSYWSHTSPEDPCFASQNRVYNDIGKEMLLHAFEGYNVCIFAYGGTGAGKSYTMM 108
          : :|| ||||| | : |: :| : :| : ||| ||||| :|||||:|
Qy   61 AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db   109 GKQESQAVIIPQLCEELFEKIND-NCNEEMSYSVEVSYMEIYCERVRDLLNPKNKGNLR 167
          ||: :||||||||||||:| : |: ||: :|||||||||||||||||||:|||
Qy   120 GYGKE-HGVI-PRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLK 177

Db   168 VREHPLLGPYVEDLSKLAVTSYTDIADLMDAGNKARTVAATNMNETSSRSHAVFTIVFTQ 227
          ||||| |||||||:||.: : : |||:|:|||||||||||||||||||||||||||: :||
Qy   178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237
```

```
Db  168 VREHPLIGPYVEDLSKLAVTSYNDIQDLMDSGNKPRTVAATNMNETSRSHAVFNIIFTQ 227
Qy  178 VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSRSHAVFTLTLTQ 237

Db  228 KRHDAETNITEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE 287
Qy  238 KWHDEETKMDTEVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db  288 MDSGPNKNKKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRY 347
Qy  298 MSSG----KQKKN-QLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRY 352

Db  348 ADRAKQIRCNAIINEDPNNKLIRELKDEVTRLRD-LLYAQGLGDITDMTNALVGMSPSS 406
Qy  353 ADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGGGGGPVEESYPPD 412

Db  407 LSALSSRAASVSSLHERILFAPGSEEAIERLKETEKIIAELNETWEEKLRRTEAIRMERE 466
Qy  413 -TPLEKQIVSIQQPDATVK-KMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEIHKERE 470

Db  467 ALLAEMGVAMREDGGTLGVFSPKKTPHLVNLNEDPLMSECLLYYIKDGVTRVGREDAERR 526
Qy  471 AALEELGISI-EKG-FVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQ 528

Db  527 QDIVLSGHFIKEEHCIFRSDSRGGGEAVVTLEPCEGADTYVNGKKVTEPSILRSGNRIIM 586
Qy  529 AEIRLNGSKILKEHCTF--EN---VDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIIL 583

Db  587 GKSHVFRFNHPEQARQERE 605
Qy  584 GDFHIFRFNHPEEARAERQ 602
```

```
RESULT   3
ENTRY        JN0114          #type complete
TITLE        kinesin-related protein unc-104 - Caenorhabditis elegans
ORGANISM                     #formal_name Caenorhabditis elegans
DATE         03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change
             29-May-1998
ACCESSIONS   JN0114
REFERENCE    JN0114
  #authors   Otsuka, A.J.; Jeyaprakash, A.; Garcia-Anoveros, J.; Tang,
             L.Z.; Fisk, G.; Hartshorne, T.; Franco, R.; Born, T.
  #journal   Neuron (1991) 6:113-122
  #title     The C. elegans unc-104 gene encodes a putative kinesin heavy
             chain-like protein.
  #cross-references MUID:91097805
  #accession JN0114
  ##molecule_type mRNA
  ##residues 1-1584 ##label OTS
  ##cross-references GB:M58582
  ##note     598-Thr and 930-Met were also found
GENETICS
  #gene      unc-104
CLASSIFICATION #superfamily kinesin-related protein unc-104; kinesin motor
             domain homology; pleckstrin repeat homology
KEYWORDS     ATP; microtubule binding; P-loop
FEATURE
  4-353      #domain kinesin motor domain homology #label KMOT\
  93-100     #region nucleotide-binding motif A (P-loop)\
  1285-1287  #region cell attachment (R-G-D) motif\
  99         #binding_site ATP (Lys) #status predicted
SUMMARY      #length 1584 #molecular_weight 179651 #checksum 5761
```

```
Query Match     36.4%;     Score 1975;  DB 1;    Length 1584;
Best Local Similarity 52.6%;   Pred. No. 0.00e+00;
Matches  318;   Conservative  123;   Mismatches 126;   Indels 38;   Gaps 20;

Db    2 SSVKVAVRVRPFNQRELSNTSKCVLQVNGNTTIN---G--HSINKE-N-F-----SFNF 49
Qy    5 GNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAFAF 64
Db   50 DHSYWSFARNDPHFITQKQVYEELGVEMLEHAFEGYNVCIFAYGQTGSGKSYTMMGKAND 109
```

```
Db  228 KKQDPETNLSTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAE 287
Qy  238 KWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297

Db  288 VSK-KKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRYADRAK 346
Qy  298 MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357

Db  347 QIKCNAVINEDPNAKLVRELKEEVTRLKDLLRAQGLGDI-IDTSMGSLTSS-PSSCSLNS 404
Qy  358 RIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGGGGGSGPVEESYPPDTPLEK 417

Db  405 QVGLTSVTSIQERIMSTPGGEEAIERLKESEKIIAELNETWEEKLRKTEAIRMEREALLA 464
Qy  418 QI-V-SIQQPDATVKKMSKAE-IVEQLNQSEKLYRDLNQTWEEKLAKTEIHKEREAALE 474

Db  465 EMGVAIREDRGDIGVFSPKKTPHLVNLNEDPLMSECLLYYIKDGITRVGQADAERRQDIV 524
Qy  475 ELGISI-E-KGFVGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIR 532

Db  525 LSGAHIKEEHCLFRSERSNTGEVIVTLEPCERSETYVNGKRVAHPVQLRSGNRIIMGKNH 584
Qy  533 LNGSKILKEHCTF--E--NVDNV-VTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFH 587

Db  585 VFRFNHPEQARAERE 599
Qy  588 IFRFNHPEEARAERQ 602
```

```
RESULT   2
ENTRY        A56921          #type complete
TITLE        kinesin family protein KIF1a - mouse
ORGANISM     #formal_name Mus musculus #common_name house mouse
DATE         02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
             13-Aug-1999
ACCESSIONS   A56921
REFERENCE    A56921
  #authors   Okada, Y.; Yamazaki, H.; Sekine-Aizawa, Y.; Hirokawa, N.
  #journal   Cell (1995) 81:769-780
  #title     The neuron-specific kinesin superfamily protein KIF1A is a
             unique monomeric motor for anterograde axonal transport of
             synaptic vesicle precursors.
  #cross-references MUID:95292344
  #accession A56921
  ##status   preliminary; translated from GB/EMBL/DDBJ
  ##molecule_type mRNA
  ##residues 1-1695 ##label RES
  ##cross-references GB:D29951; NID:g976234; PIDN:BAA06221.1; PID:g976235
CLASSIFICATION #superfamily kinesin-related protein unc-104; kinesin motor
             domain homology; pleckstrin repeat homology
KEYWORDS     P-loop
FEATURE
  6-360      #domain kinesin motor domain homology #label KMOT\
  97-104     #region nucleotide-binding motif A (P-loop)
SUMMARY      #length 1695 #molecular-weight 191723 #checksum 7412
```

```
Query Match     40.1%;     Score 2176;  DB 2;    Length 1695;
Best Local Similarity 54.0%;   Pred. No. 0.00e+00;
Matches  334;   Conservative  133;   Mismatches 121;   Indels 31;   Gaps 19;

Db    1 MAGAS-VKVAVRVRPFNSREMSRDSKCIIQMSGSTTIVNP-----K-Q-P-K---ETPK 48
Qy    1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60

Db   49 SFSFDYSYWSHTSPEDINYASQKQYRDIGEEMLQHAFEGYNVCIFAYGQTGAGKSYTMM 108
Qy   61 AFAFDRSYWSFDK-NAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db  109 GKQEKDQQGIIPQLCEDLFSRINDTTND-NMSYSVEVSYMEIYCERVRDLLNPKNKGNLR 167
Qy  120 G-YGKEH-GVIPRICQDMFRRINELQKDKNLTCTVEVSVLEIYNERVRDLLNPSTKGNLK 177
```

```
Qy   65 DRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG--YG 122
         |:|||| :|:|||  |  |:::||| || |  |||||||||||||||||||||||||  ||
Db  110 PDEMGIIPRLCNDLFARIDN-NNDKDVQYSVEVSYMEIYCERVKDLLNPSGGNLRVREH 168

Qy  123 KEH-GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181
         |:| ||||| ||| :||:  |:  ::|||: |||||:|||:|||||||||: ||||:|||
Db  169 PLLGPYVDDLTKMAVCSYHDICNLMDEGNKARTVAATNMNSTSSRSHAVFTIVLTQKRHC 228

Qy  182 PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241
         |  |||::||:|:  :| |:| |||||||||||||||||| ||||||||||:| ||| |
Db  229 ADSNLDTEKHSKISLVDLAGSERANSTGAEGQRLKEGANINKSLTTLGLVISKLAEESTK 288

Qy  242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301
         :::|| |||:||||||||||||| |||||:|||||||| |:|||:|||:|||  |||:||
Db  289 KKKSNKGVIPYRDSVLTWLLRENLGGNSKTAMLAALSPADINFDETLSTLRYADRAKQIV 348

Qy  302 KQKKNQ-LVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK 360
         :  |:| :||||||||||||::|||||| |||:||||||||||:|||||||||| ||:|
Db  349 CQAVVNEDPNAKLIRELNEEVIKLRHILKDKG-----IDVTD--VQET--PGK--HKK-G 396

Qy  361 NHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPDTPLEKQIV 420
         :|||||||||:|||||:||:| :::||       |      :  :  |:||:||||||
Db  397 P-KLP-AHVH----EQL-EKLQESEKLMAEIGKTWEQKLIHTEIRKQREEELRDMGLAC 449

Qy  421 SIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEIHKEREALEELGIS- 479
            : ||||||||:|||:||||:||:||||| ||:|||||||||||:|||||||||:
Db  450 AEDGTTLGVFSPKKLPHLVNLNEDPLMSECLIYYLKEGVTSVGRPAEHRPDILLSGEAI 509

Qy  480 IEKGF-VGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIRLNGSKI 538
          |:|  ||||:|||:|||||| |||| |||: |:||||||||||:|:||:||||:|||||
Db  510 LELHCEFINEDGNVTLTMKPNASCYINGKQVTTPTVLHTGSRVILGEHHVFRYNDPQEAR 569

Qy  539 LKEHCTFENVDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEAR 598
          |:| :|||:|||:| :| ::  |:|||||||||||||||||||||||||||||:|:|||
Db  570 QSRHN 574

Qy  599 AERQE 603
         |:|
```

RESULT 4
ENTRY   T15822   #type complete
TITLE   kinesin-like protein unc-104 - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE   20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
          20-Sep-1999
ACCESSIONS T15822
REFERENCE Z18412
#authors   Du, Z.
#submission submitted to the EMBL Data Library, February 1996
#description The sequence of C. elegans cosmid C52E12.
#accession T15822
##status   preliminary; translated from GB/EMBL/DDBJ
##molecule_type DNA
##residues 1-1584 ##label DU2
##cross-references EMBL:U50135; NID:g1208884; PID:g1208885;
                   PIDN:AAA93453.1; CESP:unc-104
GENETICS CESP:unc-104
#gene   unc-104
#introns   34/1; 84/2; 140/3; 200/2; 285/3; 339/2; 490/2; 580/1; 631/3;
           672/1; 738/2; 782/2; 856/3; 902/1; 942/2; 991/3; 1045/2;
           1256/3; 1352/3; 1397/1; 1545/2
SUMMARY #length 1584 #molecular-weight 179649 #checksum 6627

Query Match 36.4%; Score 1971; DB 2; Length 1584;
Best Local Similarity 52.6%; Pred. No. 0.00e+00;
Matches 318; Conservative 122; Mismatches 127; Indels 38; Gaps 20;

```
Db    2 SSVKVAVRVRPFNQREISNTSKCVLQVNGNTTTIN---G--HSINKE-N-F-----SFNF 49
          : |:|||||||||:|||  :::|:  |::| :|| |   |  :|| | | |     |: |
Qy    5 GNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKAFAF 64

Db   50 DHSYWSFARNDPHFITQKQVYEELGVEMLEHAFEGYNVCIFAYGQTGSGKSYTMMGKAND 109
         |:||||| :|:|| |||:|| :||:||:|  || |||||||||||||||||:||: |:||
```

```
Qy   65 DRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG--YG 122
         |:|||| :|:|||  |  |:::||| || |  |||||||||||||||||||||||||  ||
Db  110 PDEMGIIPRLCNDLFARIDN-NNDKDVQYSVEVSYMEIYCERVKDLLNPSGGNLRVREH 168

Qy  123 KEH-GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181
         |:| ||||| ||| :||:  |:  ::|||: |||||:|||:|||||||||: ||||:|||
Db  169 PLLGPYVDDLTKMAVCSYHDICNLMDEGNKARTVAATNMNSTSSRSHAVFTIVLTQKRHC 228

Qy  182 PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241
         |  |||::||:|:  :| |:| |||||||||||||||||| ||||||||||:| ||| |
Db  229 ADSNLDTEKHSKISLVDLAGSERANSTGAEGQRLKEGANINKSLTTLGLVISKLAEESTK 288

Qy  242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301
         :::|| |||:||||||||||||| |||||:|||||||| |:|||:|||:|||  |||:||
Db  289 KKKSNKGVIPYRDSVLTWLLRENLGGNSKTAMLAALSPADINFDETLSTLRYADRAKQIV 348

Qy  302 KQKKNQ-LVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIK 360
         :  |:| :||||||||||||::|||||| |||:||||||||||:|||||||||| ||:|
Db  349 CQAVVNEDPNAKLIRELNEEVIKLRHILKDKG-----IDVTD--VQET--PGK--HKK-G 396

Qy  361 NHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPDTPLEKQIV 420
         :|||||||||:|||||:||:| :::||       |      :  :  |:||:||||||
Db  397 P-KLP-AHVH----EQL-EKLQESEKLMAEIGKTWEQKLIHTEIRKQREEELRDMGLAC 449

Qy  421 SIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEIHKEREALEELGIS- 479
            : ||||||||:|||:||||:||:||||| ||:|||||||||||:|||||||||:
Db  450 AEDGTTLGVFSPKKLPHLVNLNEDPLMSECLIYYLKEGVTSVGRPVAEHRPDILLSGEAI 509

Qy  480 IEKGF-VGPYHSKEMPHLVNLSDDPLLAECLVYNIKPGQTRVGNVNQDTQAEIRLNGSKI 538
          |:|  ||||:|||:|||||| |||| |||: |:||||||||||:|:||:||||:|||||
Db  510 LELHCEFINEDGNVTLTMKPNASCYINGKQVTTPTVLHTGSRVILGEHHVFRYNDPQEAR 569

Qy  539 LKEHCTFENVDNVVTIVPNEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHPEEAR 598
          |:| :|||:|||:| :| ::  |:|||||||||||||||||||||||||||||:|:|||
Db  570 QSRHN 574

Qy  599 AERQE 603
         |:|
```

RESULT 5
ENTRY   T13827   #type complete
TITLE   kinesin-73 - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE   20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
          20-Sep-1999
ACCESSIONS T13827
REFERENCE Z17784
#authors   Li, H.P.; Liu, Z.M.; Nirenberg, M.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1997) 94:1086-1091
#title   Kinesin-73 in the nervous system of Drosophila embryos.
#accession T13827
##status   preliminary; translated from GB/EMBL/DDBJ
##molecule_type mRNA
##residues 1-1921 ##label LIH
##cross-references EMBL:U81788; NID:g1906595; PID:g1906596;
                   PIDN:AAB50404.1
SUMMARY #length 1921 #molecular-weight 215047 #checksum 9262

Query Match 29.5%; Score 1597; DB 2; Length 1921;
Best Local Similarity 59.6%; Pred. No 1.82e-262;
Matches 238; Conservative 66; Mismatches 84; Indels 11; Gaps 8;

```
Db    1 MASD-KIKVAVRVRPFNRREIELDTKCIVEMEKQQTILQNPP-PLEKIER-K--Q-PK 52
         | :: :|||:|||||||  |||:  ||||:||  |||  |   |  | || |  | ||
Qy    1 MSGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPK 60

Db   53 TPAFDHCFYSLNPEDENFASQETVFDCVGRGILDNAFQGYNACIFAYGQTGSGKSYTMMG 112
          :|||  :|| ||| |:|||  |:|| || | |||||:||| ||||||||||||||:|||
Qy   61 AFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMG 120

Db  113 TQESKGIIPRLCDQLFSAIAN-KSTPELMYKVEVSYMEIYNEKVHDLLDPKPNKQSLKVR 171
         | :::|||||||:  | :| | |: || |:||||||||||||:| |||  :||:|||
```

```
Qy   121  YGKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNP-STKGNLKVR 179
          |||||||||||||||||||||||||| :||||||:||||||||||||||| :||||||||
Db   172  EHNVMGPYVDGLSQLAVTSYQDIDNLMTEGNKSRTVAATNMNAESSRSHAVFSVVLTQIL 231
          || :|| |:  |  |  :  ::| || |:||| ||||||||| ||||:||| ::::| :
Qy   180  EHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKW 239
          ::  || |:  :  : |  :: | :: :::|| || ||| | :|| ::|:|  | : :
Db   232  TDQATGVSGEKVSRMSLVDLAGSERAVKTGAVGDRLKEGSNINKSLTTLGLVISKLADQS 291
          :        || :||||||||||||   :|| :|:||||:  |:||:|||||||| |:|
Qy   240  HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMS 299
Db   292  NGKKSGNDKFVPYRDSVLTWLLKDNLGGNSRTVMATISPSADNYEETLSTLRYADRAKR  351
          :   :::  |||||||||||||||:||||| || | ||| || : ||||||||||| |||
Qy   300  SGKQK-KNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKR 358
          ::     |||||||||||||||||:|||||  :|| |||| || ||||||||||||| |||
Db   352  IVNHAVVNEDPNARIIRELRHEVETLRSMLKHATGSPVG               390
          |  ||||||||||| ||||: |  ||||::|
Qy   359  IKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGG               397
```

RESULT   6

ENTRY   T16759   #type complete

TITLE   hypothetical protein R144.1 - Caenorhabditis elegans

ORGANISM   #formal_name Caenorhabditis elegans

DATE   20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
        20-Sep-1999

ACCESSIONS   Z18571

REFERENCE   T16759
    #authors   Favello, T.
    #submission   submitted to the EMBL Data Library, March 1995
    #description   The sequence of C. elegans cosmid R144.
    #accession   T16759
    ##status   preliminary; translated from GB/EMBL/DDBJ
    ##molecule_type DNA
    ##residues   1-928 ##label FAV
    ##cross-references EMBL:U23515; NID:g746492; CESP:R144.1;
                      PIDN:AAC46545.1; CESP:R144.1 PID:g746493;

GENETICS
    #gene   CESP:R144.1
    #introns   80/3; 125/1; 261/2; 331/3; 350/2; 516/3; 589/1; 615/2; 638/3;
                726/1; 865/3; 901/2

SUMMARY   #length 928 #molecular-weight 103589 #checksum 328

Query Match   23.8%;
Best Local Similarity   49.8%;   Score 1291;   DB 2;   Length 928;
Matches   199;   Conservative   81;   Pred. No. 3.6le-205;
Mismatches 113;   Indels 7;   Gaps 6;

```
Db    1  MGKGDSIIVAVRVRPFNDREKTRNCKLVIEMPDEETTVIRDPKTNDEKRFTYDHSYW-SH 59
         |.|||..|.|||||||.||.|.|.||.|.||..|.:|.||.:|.::||.|.:|:.:| .|
Qy    1  MSGGGNIKVVVRVRPFNAREIDRGAKCIVRM-EGNQTILTPPPGAEEKARKSGKTIMDGP 59
Db   60  DGFS-EKKNGYLEPTDPHYADQRRVFEDLGRGVLANAWAGYNCSLFAYGQTGSGKSYSIV 118
         .||: .: ||.|.:.|::|||||||.||||.|||.||..||||.||||||||||||||.:
Qy   60  KAFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119
Db  119  GFKNNKGIVPIVCEELFKQIAD-NKKKNMQFEVFVSMEIYCEKVRDLLSSTPPPKGGLK  177
         |:...:|:.|...:::|.:|..  ::.|..:..|.|.||.||||||||.:. |..||.||
Qy  120  GYGKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLN--PSTKGNLK 177
Db  178  VREHPKNGFVENLTTVPVNSFKEIEAKIEEGTKSRTIAATQMNATSSRAHTIVKITFNQ  237
         |||||..|..|.|:...|.|||:|....:..||.||.|||..||:|||:.|:..::..|
Qy  178  VREHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQ 237
Db  238  KSSKQAGGTSMKK-SEINLVDLAGSERQSAAGTEGDRLKEGIVINQSLTTLGRVIKALHD 296
         |...:...||.:: |.|.:||||||||::|:.|:..||||||  |:||:|||||||.||.
Qy  238  KWHIDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD 297
Db  297  SQAKSGKKTQIPYRDSVLTCLLKNALGNGSKTIMIAAISPADINFEETLSTLRFADRAK  356
         .......||..|||||||| |||:.||.|| .|.|||||||||||||||||||||.||.||
Qy  298  MSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAK 357
Db  357  SIKTNAVVNENQTERALRELREENLRLQSOIQGGTAGDAS               396
```

```
Qy  358  RIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGG               397
```

RESULT   7

ENTRY   S38982   #type complete

TITLE   kinesin-related protein KRP85 - sea urchin
         (Strongylocentrotus purpuratus)

ALTERNATE_NAMES   kinesin-2 chain A; KRP (85/95) 85K chain

ORGANISM   #formal_name Strongylocentrotus purpuratus #common_name
            purple urchin

DATE   10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
        10-Sep-1999

ACCESSIONS   S38982; S72551

REFERENCE   S38982
    #authors   Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.;
                Scholey, J.M.
    #journal   Nature (1993) 366:268-270
    #title   Novel heterotrimeric kinesin-related protein purified from
              sea urchin eggs.
    #cross-references MUID:94050179
    #accession S38982
    ##molecule_type mRNA
    ##residues   1-699 ##label COL1
    ##cross-references EMBL:L16993; NID:g295245; PIDN:AAA16098.1;
                      PID:g295246
    #accession S72551
    ##molecule_type protein
    ##residues   2-5,'X',7-11;59-64;125-132;222-226,'X',228-230 ##label
                  COL2

COMPLEX   heterotrimer of a 115K chain and two kinesin-related chains
           of 95K (PIR:S58691) and 85K

CLASSIFICATION   #superfamily kinesin-related protein KIF3; kinesin motor
                  domain homology

KEYWORDS   ATP; heterotrimer; microtubule binding; P-loop

FEATURE
    11-348   #domain kinesin motor domain homology #label KMOT\
    97-104   #region nucleotide-binding motif A (P-loop)\
    103   #binding_site ATP (Lys) #status predicted

SUMMARY   #length 699 #molecular-weight 78697 #checksum 1222

Query Match   18.6%;
Best Local Similarity   45.0%;   Score 1008;   DB 1;   Length 699;
Matches   182;   Conservative   86;   Pred. No. 1.1e-152;
Mismatches 111;   Indels 25;   Gaps 15;

```
Db    6  SGDNVRVVVRCRPLNSKETGQGFKSVVKMDEMRG--TVQV-TNPNAP-SG----EPPKS 57
         || :..|||| ||:|:.:.  ||.|| :|   .   |.:. ||..|  .|    |..:
Qy    2  SGGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKA 61
Db   58  FTFD-T--VF---APG-AKQTDVYNQTARPIVDAIIEGYNGTIFAYGQTGTGKTFTMEG- 109
         |:|| :  ||   |.| |:|.:.  ::.  .:::...:||...||||||||:||::.:|
Qy   62  FAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY 121
Db  110  VRSQPELRGIIPNSFAHIFGHIAKEQENVRFLVRVSYLEIYNEEVKDLLGKDQQHRLEVK 169
         :::||.|.||..:.|..|..|:.||.||.. ..||||||||||.|:|||   :...|.|:
Qy  122  GKEHGVIPRICQDMFRRI-NELQKDK-NLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVR 179
Db  170  ERPDVGVYVKDLSAFVVNNADDMDRIMTLGNKNRSVGATNMNESSSRSHAIFITL-ERS  228
         |.|..|.||:||:  ||..:.::.:.|..||| || |||||||:||||||| |.:| ::
Qy  180  EHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKW 239
Db  229  DMGLDKEQHVRVGKLHMVDLAGSERQKTKGTQGQRLKEATKINLSLSTLGNVISSLVD-- 286
         .:.:||:  |.| ||.||||||||||:|:.||.|||||..:||.||||||.||::|.
Qy  240  HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMS 299
Db  287  -GK--STH-IPYRNSKLTRLLQDSLGGNAKTVMCANIGPAEYNYDETISTLRYANRAKNI 342
          ||  ||. |||| ||||||||:|||||: .||  .|.|||.|||||:||||||:| |:|
Qy  300  SGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRI 359
Db  343  KNKAKINEDPKDALLREFQKEIEELKQISESGEGLDDDEESGS               386
         ||:|.|||||:.|::||:|:||.|||||:.|.|| |.|.|||.
Qy  360  KNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGG               403
```

Qy   422 IQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALE 474

RESULT   9
ENTRY   S58691   #type complete
TITLE   kinesin-related protein KRP95 - sea urchin
              (Strongylocentrotus droebechiensis)
ALTERNATE_NAMES kinesin-2 chain B; KRP (85/95) 95K chain
ORGANISM   #formal_name Strongylocentrotus droebechiensis
DATE   10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
              10-Sep-1999
ACCESSIONS   S58691
REFERENCE   S58691
#authors   Rashid, D.J.; Wedaman, K.P.; Scholey, J.M.
#journal   J. Mol. Biol. (1995) 252:157-162
#title   Heterodimerization of the two motor subunits of the
              heterotrimeric kinesin, KRP(85/95).
#cross-references MUID:95404610
#accession   S58691
##status   preliminary; nucleic acid sequence not shown; not
                      compared with conceptual translation
##molecule_type mRNA
##residues   1-742 ##label RAS
COMPLEX   heterotrimer of a 115K chain and two kinesin-related chains
              of 85K (PIR:S39982) and 95K
CLASSIFICATION #superfamily kinesin-related protein KIF3; kinesin motor
              domain homology
KEYWORDS   ATP; heterotrimer; microtubule binding; P-loop
FEATURE
  9-345   #domain kinesin motor domain homology #label KMOT\
  95-102   #region nucleotide-binding motif A (P-loop)\
  101   #binding_site ATP (Lys) #status predicted
SUMMARY   #length 742 #molecular-weight 84156 #checksum 2976

Query Match   17.7%;   Score 962;   DB 1;   Length 742;
Best Local Similarity 44.5%;   Pred. No. 3.24e-144;
Matches   177; Conservative 86; Mismatches 105; Indels 30;   Gaps 18;

Db   5 SAETVKVVVRCRPMNSKEISQGHKRIVEMDNKRG-LVEVTNP--KGP-PGEP--N--KSF 56
           :: |||||| ||:|||||| :|:|| ::| :| ::| |: :: |||:|:
Qy   3 GGGNIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAF 62

Db   57 TFDTVVDWN-SKQIDLY---DETFRSL-V---ESVLQGFNGTIFAYGQTGKTFTMEGV 108
           :|| |: | ::| | : |: :|: || ::::||||||||| ::::|||
Qy   63 AFDRSY-WSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY 121

Db   109 RSNPELRGVIPNSFEHIFTHIARTQNQQFL---VRASYLEIYQEEIRDLLAKDQKKRLDL 165
           : :|||||||||||| ||| | ||||| |||| | ||||
Qy   122 -GK-E-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKV 178

Db   166 KERPDTGVYYKDLSSFVTKSVKEIEHVMTVGNNNRSVGSTNMNEHSSRSHAIFIITIECS 225
           ::| ||| || ||: :| ||||: ::::||| ||||| :|||::||:||:||||| ::|:
Qy   179 REHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQK 238

Db   226 ELGVDGENHI-RVGKLNLVDLAGSERQAKTGATGDRLKEATKINLSLSALGNVISALVD- 283
           ||| || ::| :||| |:|||||||| ||| ::||||||:|||::||:|||:::| |
Qy   239 WHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADM 298

Db   284 --GK--SSH-IPYRDSKLTRLLQDSLGGNAKTVMVANMGPASYNPDETITTLRYANRAKN 338
           || ::| |||||||||||||||| ||||:::||: :|:|| || ::|::|||||:|
Qy   299 SSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKR 358

Db   339 IKNKPKINEDPKDALLREFQEEISRLKQALDKKGPSDG 376
           |||: :|||| || ||||||||||||| | |||
Qy   359 IKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGG 396

RESULT   10
ENTRY   A47334   #type fragment
TITLE   Lckin kinesin-related antigen - Leishmania chagasi (fragment)
ORGANISM   #formal_name Leishmania chagasi
DATE   21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
             24-Sep-1999

RESULT   8
ENTRY   B44259   #type complete
TITLE   kinesin-related protein KIF3A - mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE   03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change
              22-Jun-1999
ACCESSIONS   B44259; S27872
REFERENCE   A44259
#authors   Aizawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.;
              Hirokawa, N.
#journal   J. Cell Biol. (1992) 119:1287-1296
#title   Kinesin family in murine central nervous system.
#cross-references MUID:93077686
#accession   B44259
##molecule_type mRNA
##residues   1-701 ##label AIZ
##cross-references EMBL:D12645; NID:g220469; PIDN:BAA02166.1;
              PID:d1002656; PID:g220470
##experimental_source brain
##note   sequence extracted from NCBI backbone (NCBIP:118911)
COMPLEX   heterodimer with KIF3B (PIR:A57107); the KIF3A/3B heterodimer
              associates with kinesin superfamily-associated protein
              (KAP3) (PIR:JC6161) to form a heterotrimer
FUNCTION
#description KIF3 complex is a motor protein that provides anterograde
              fast axonal transport of membranous organelles
CLASSIFICATION #superfamily kinesin-related protein KIF3; kinesin motor
              domain homology
KEYWORDS   ATP; coiled coil; heterodimer; heterotrimer; microtubule
              binding; P-loop
FEATURE
  1-368   #domain head globular #status predicted #label HGL\
  15-351   #domain kinesin motor domain homology #label KMOT\
  100-107   #region nucleotide-binding motif A (P-loop)\
  369-599   #domain helical rod #status predicted #label ROD\
  600-701   #domain tail globular #status predicted #label TGL\
  106   #binding_site ATP (Lys) #status predicted
SUMMARY   #length 701 #molecular-weight 80167 #checksum 60

Query Match   17.7%;   Score 962;   DB 1;   Length 701;
Best Local Similarity 41.4%;   Pred. No. 3.24e-144;
Matches   196; Conservative 104; Mismatches 137; Indels 36;   Gaps 28;

Db   14 NVKVVVRCRPLNERE--K-SMCY-R-QAV-SV-DEMRGTI-TVHKTDSSN-EPPKTFTFD 64
           :|||||| || | ||: || :| |::| ::::|| | |:| ::|: ||| | |||
Qy   6 NIKVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPGAEEKARKSGKTIMDGPKAFAFD 65

Db   65 -T--VF---GPE-SKQLDVY-NLTARPIIDSVLEGVNGTIFAYGQTGKTFTMEGVRAV 116
           | :| | ||||| | | :|:|| |:|||||||| :::|||
Qy   66 RSYWSFDKNAPNYARQEDLFQDLGV-PLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-G- 122

Db   117 PGLRGVIPNSFAHIFGHIAKAEGDTRFL--VRVSYLEIYNEEVRDLLGKDQTQRLEVKER 174
           || ::|| :|: :| ::| ::| :| : :| ||||||:| ||||| :||:: ||:
Qy   123 KE-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181

Db   175 PDVGVVIKDLSAYVVNNADDMDRIMTLGHKNRSVGATNMNEHSSRSHAIFTITIECSEKG 234
           ||| :: ||:| :||: :| :|| : | |:||:||:|||||:||:||:||| ::|:
Qy   182 PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241

Db   235 VDGNMHV-RMGKLHLVDLAGSERQAKTGATGQRLKEATKINLSLSTLGNVISALVD---G 290
           | ::| : :||:|:||||||||:||| ::| |||||:|||::| ||||:::||| | ||
Qy   242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301

Db   291 K--STH-VPYRNSKLTRLLQDSLGGNSKTMMCANIGPADYNYDETISTLRYANRAKNIKN 347
           | :|| |||||||||| |||||:|:|| : ||:|||:|: :|:| ||||:|:|||
Qy   302 KQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKN 361

Db   348 KARINEDPKDALLRQFQKEIEELKKKLEE-GEEVSGSDISGSEEDDE-EGELGEDGEKRK 405
           |: ||||| || |:|| ::|:| :| :: : |::| | | | |:|| | | |:|:|:|
Qy   362 HAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGGSGGGPGGGGSGGGPVEESYPPDTPLEKQIVS 421

Db   406 KRRDQAGKKKVSPDKMVE-M-QA-KIDEE-RKALETKLDMEEEERNKARAELE 454
           :: ||:| |||| :|| :: || |::|| ||:|||| :

ACCESSIONS   A47334
REFERENCE   A47334
  #authors   Burns Jr., J.M.; Shreffler, W.G.; Benson, D.R.; Ghalib, H.W.;
                Badaro, R.; Reed, S.G.
  #journal   Proc. Natl. Acad. Sci. U.S.A. (1993) 90:775-779
  #title   Molecular characterization of a kinesin-related antigen of
                Leishmania chagasi that detects specific antibody in
                African and American visceral leishmaniasis.
  #cross-references MUID:93133867
  #accession   A47334
    ##status   preliminary
    ##molecule_type DNA
    ##residues 1-955 ##label BUR
    ##cross-references GB:L07879; NID:g308884; PIDN:AAA29254.1; PID:g308885
    ##experimental_source MHOM/BR/82/BA-2.C1
    ##note   sequence extracted from NCBI backbone (NCBIN:122864,
                NCBIP:122865)
CLASSIFICATION #superfamily unassigned kinesin-related proteins; kinesin
                motor domain homology
KEYWORDS   ATP; P-loop
FEATURE
     13-398   #domain kinesin motor domain homology #label KMOT\
    122-129   #region nucleotide-binding motif A (P-loop)
SUMMARY   #length 955 #checksum 7462

Query Match   17.6%; Score 953; DB 2; Length 955;
Best Local Similarity   43.5%;
  Pred. No. 1.46e-142;
Matches 177; Conservative   90; Mismatches 111; Indels   29; Gaps   18;

Db   13 VKVSVRVRPLNERENNAPEGTKVTVAAKQAAAVTVKVLGGSNNSGAAESMGTARRVAQD 72
            : | ||||||||| : : ||| : :: || :: | :: ::||
Qy   7 IKVVVRVRPFNAREID-R-GAKCIVR-MEGNQTI-LTPPPGAEEKARKSGK-TIMDGPKA 61

Db   73 FQFDHVFWSVETPDACGATPATQADVFRTIGYPLVQHAFDGFNSCLFAYGQTGSGKTYTM 132
          | || :| :: :|| : ||| : |||:|||||||||||:|
Qy   62 FAFDRSYWSFDK-NA-PNY-ARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSM 118

Db   133 MGADVSALSGEGNGVTPRICLEIFARKASVEAAQGHSRWIVELGVYEVYNERVSDLLGKRK 192
          || : :: ||||: ::| : : | | ||:| |:|:|||||| ||| ::
Qy   119 MG-----Y-GKEHGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRDLL-NP- 170

Db   193 KGVKGGGEEVYVDVREHPSRGVFLEG-QRLVEVGSLDDVVRLIEIGNGVRHTASTKMNDR 251
          : || ||||||||: | ::| ||||: :: |||:|||||| ||:|||||||:
Qy   171 -STKG---NL--KVREHPSTGPYVEDLAKLV-VRSFQEIENLMDEGNKARTVAATNMNET 223

Db   252 SSRSHAIIMLLLREERTMTTKSGETIRTAGKSSRMNLVDLAGSERVAQSOVEGGQQFKEAT 311
          |||||| :|| : : :: | || | || ::|||||||||: | ::|| ||
Qy   224 SSRSHAVFTLTLTQKW----HDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGA 279

Db   312 HINLSLTTLGRVIDVLADMATKGAKAQYSVAPFRDSKLTFILKDSLGGNSKTFMIATVSP 371
          |||||| |||||| |||| : | : | ||| ||:||:||||||||||: | ||| ||
Qy   280 EINRSLSTLGRVIAALADMSS-GKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISP 338

Db   372 SALNEYEETLSTLRYASRARDIVNVAQVNEDPRARRIRELEEQMEDMR 418
            :| |:|||||||||:| | ||:|:| ||||||:|||: |: |:|:
Qy   339 ADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLR 385

RESULT 11
ENTRY   A57107   #type complete
TITLE   kinesin-related protein KIF3B - mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE   03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change
                22-Jun-1999
ACCESSIONS   A57107
REFERENCE   A57107
  #authors   Yamazaki, H.; Nakata, T.; Okada, Y.; Hirokawa, N.
  #journal   J. Cell Biol. (1995) 130:1387-1399
  #title   KIF3A/B: a heterodimeric kinesin superfamily protein that
                works as a microtubule plus end-directed motor for membrane
                organelle transport.
  #cross-references MUID:96032268

    ##status   nucleic acid sequence not shown
    ##molecule_type mRNA
    ##residues 1-747 ##label YAM
    ##cross-references GB:D26077; NID:g1060922; PIDN:BAA05070.1;
                PID:d1005611; PID:g1060923
    ##experimental_source brain
COMPLEX   heterodimer with KIF3A (PIR:B4259); the KIF3A/3B heterodimer
                associates with a kinesin superfamily-associated protein
                (KAP3) (PIR:JC6161 in testis) to form a heterotrimer
FUNCTION
  #description KIF3 complex is a motor protein that provides anterograde
                fast axonal transport for an unknown cargo
CLASSIFICATION #superfamily kinesin-related protein KIF3; kinesin motor
                domain homology
KEYWORDS   ATP; coiled coil; heterodimer; heterotrimer; microtubule
                binding; P-loop
FEATURE
      1-363   #domain head globular #status predicted #label HGL\
     10-346   #domain kinesin motor domain homology #label KMOT\
     96-103   #region nucleotide-binding motif A (P-loop)\
    364-592   #domain helical rod #status predicted #label ROD\
    594-747   #domain tail globular #status predicted #label TGL\
    102   #binding_site ATP (lys) #status predicted
SUMMARY   #length 747 #molecular_weight 85288 #checksum 7951

Query Match   17.4%; Score 942; DB 1; Length 747;
Best Local Similarity   46.7%; Pred. No. 1.53e-140;
Matches 163; Conservative   73; Mismatches   99; Indels   14; Gaps   9;

Db   37 KLGQVSVKNPKGTSHEMPKTFTFDAVYDWNAKQFELYDETFRPLVDSVLQGFNGTIFAYG 96
          | : : || | | :| | :: ::   ::| :: | ::| : ::|||||
Qy   50 KSGKTIMDGPKAFAFDRS-YWSFDKNAPNYAROEDLFQDLGVPLLDNAFKGYNNCIFAYG 108

Db   97 QTGTGKYTMEGVRGDPEKRGVIPNSFDHIFTHISRSQNQQY---VRASYLEIYQEEIR 153
          |||:|| : : : | ||. |:| :| ::   .:|||||: || :
Qy   109 QTGSGKSYSMSMGY-GK-E-HGVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVR 165

Db   154 DLLSKDQTKRLELKERPDTGVVYVKDLSSFVTKSVKEIEHVMNVGNQNRSVGATNMNEHSS 213
          ||| :   : : || || ||:| |:: |:||::|| :| :||:| || |||||:||
Qy   166 DLLNPSTKGNLKVREHPSTGPYVEDLAKLVVRSFQEINLMDEGNKARTVAATNMNETSS 225

Db   214 RSHAIFVTTI-ECSEVGLDGENHIRVGKLNLVDLAGSERQAKTGAQGERLKEATKINLSL 272
          ||||:| | | : :: :   : : |:::|:|||||||| ||| :|||||| :||.|
Qy   226 RSHAVFTLTLTQKWHDEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSL 285

Db   273 SALGNVISALVD--GK--STH-IPYRDSKLTRLLQDSLGGNAKTVMVANVGPASYNVEE 326
          | ||.|::||.| || | :|||||:||::|:|||||||: | : |: || :| :|
Qy   286 STLGRVIAALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEE 345

Db   327 TLTTLRYANRAKNIKNKPRVNEDPKDALLREFQEEIARLKAQLEKRSIG 375
          || ||||| .||.|||: :||||| .|::|| :||:||:| |:|::
Qy   346 TLSTLRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGG 394

RESULT 12
ENTRY   S38983   #type fragment
TITLE   kinesin-related protein 95K chain - sea urchin
                (Strongylocentrotus purpuratus) (fragment)
ORGANISM   #formal_name Strongylocentrotus purpuratus #common_name
                purple urchin
DATE   07-Oct-1994 #sequence_revision 07-Feb-1997 #text_change
                10-Jul-1998
ACCESSIONS   S38983; S72552
REFERENCE   S38982
  #authors   Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.;
                Scholey, J.M.
  #journal   Nature (1993) 366:268-270
  #title   Novel heterotrimeric kinesin-related protein purified from
                sea urchin eggs.
  #cross-references MUID:94050179
  #accession   S38983
    ##status   nucleic acid sequence not shown
    ##molecule_type mRNA

```
11-359        #domain kinesin motor domain homology #label KMOT\
97-104        #region nucleotide-binding motif A (P-loop)
SUMMARY       #length 786 #molecular-weight 86671 #checksum 3764

Query Match  16.0%;  Score 870;  DB 2;  Length 786;
Best Local Similarity 40.0%;  Pred. No. 2.37e-127;
Matches  187;  Conservative 105;  Mismatches 138;  Indels 38;  Gaps 22;
```

```
Db    6  GGSESVKVVRCRPLNGKEKADGRSRIVDMD----V-DA--G-QVKVRNPKADASEPPKA 57
         :|:..:||||:||:|:|                : |. .  |  :.  ||: |  : ||
Qy    2  SGGGNIKVVVVRVRPFNAREIDRGAKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGPKA 61

Db   58  FTFDQVY--DWNCQQ--R--DVFDITARPLIDSCIEGYNGTIFAYGQTGTGKSHTMEG- 109
         |.||:.|  |.|.    .    |||:.:| : |  | |||| |||||||||:|||:. |
Qy   62  FAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY 121

Db  110  KDEPPELRGLIPNTFRYVFEIIARDSGTKEFLVRSSYLEIYNEEVRDLLGKDHSKKMELK 169
         |.|  ..  :||: |: :.|:  .     :  :: :|||||||:||||||.  .: :||  .
Qy  122  GKEHGVIPRICQDMFRRINEL-QKDKNLT-CTVEVSYLEIYNERVRDLLNPSTKGNLKVR 179

Db  170  ESPDRGVYVKDLSQFVCKNYEEMNKVLLAGKDNRQVGATLMNQDSSRSHSIFTITIECIE 229
         | |  ||||:||:.:| : .:|:|: :.  |.|. :. | |:| ||||||::|:| . | .
Qy  180  EHPSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTL--TQ 237

Db  230  KLESAAAQKPGAKKDDSNHVRVGKLNLVDLAGSERQDKTGATGDRLKEGIKINLSLTALG 289
         |  .    |.|.||..     | |.:.|||||||| :::||||| |||| :|| ||:|:||
Qy  238  KWHDE-E---TKMD-TE--KVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLG 289

Db  290  NVISALVD---GK--SGH-IPYRDSKLTRLLQDSLGGNTKTVMVANIGPADWNYDETMST 343
         .|::||.|   ||  ..  :|||||:|| |||:||||||::| :|. | || |: || |
Qy  290  RVIAALADMSSGKQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLST 349

Db  344  LRYANRAKNIQNKPKINEDPKDAMLRQFQEEIKIKEQLAARAAGGGGPI-TM-PSGGGS 401
         ||||:.||.|:|. .:|||| :.|:|:::||.|::|.|  : ||||   :   :| ...
Qy  350  LRYADSAKRIKNHAVVNEDPNARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESY 409

Db  402  PTQKIVER-TEEVDPDIDAIKAQMRAELEAKMK-SDISTEALDKAREE 447
         |.:| : . ..: .   .| :.::.:::  : | ::|..: .  :.::
Qy  410  PPDTPLEKQIVSIQQPDATVKKMSKAEIVEQLNQSEKLYRDLNQTWEE 457
```

```
Db  170  -FFSYDQVYYSCDPERKNFACQAKVFEGTARPLIDTAFEGYNACLFAYGQTGSGKSYSMM 228
          |:.:.| : : .    :|. ||.:::    | |:|.||:|||.|:|||||||||||||||
Qy   60  KAFAFDRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMM 119

Db  229  GIEALDDAALDGGPHDEAGIIPRFCHELFRRIEAVKSQQQLQVEVESYFEIYNEKIHD 288
         | ..|  .      .|   .|:||.::::||||.:.::|.:.  ||:|:|||||:: .|
Qy  120  GY-G--K---E----H--GVIPRICQDMFRRINELQKDKNLTCTVEVSYLEIYNERVRD 166

Db  289  LLSVQHAAAATGESTPIQQQQQQRPALKVREHPIFGPYVVDLSAHSVDSYSALRNWLAV 348
         ||                      .:LKVREHPS.GPYV.DL|.|.V:S.|.|..::.:
Qy  167  LL---NPST-KGN----------LKVREHPSTGPYVEDLAKLVVRSFQEIENLMDE 208

Db  349  GNSQRATASTAMNDKSSRSHSIFNIVLNLTDLSSDDGLSSDTDSSTASSLRQTRRSKISL 408
         |||:|:|.||:.|:.|||||:.|.:.:.:.. .. .:  .|..  .: .  :. .||||
Qy  209  GNKARTVAATNMNETSSRSHAVFT--LTLTQ-KWHD----E-ETKMDTE-K-V--AKISL 256

Db  409  VDLAGSERISVSGSNGERIREGVSINKSLLTLGKVIAALADSRKASANGPLGSGTPSTFV 468
         ||||||||.:.:|:.|.|::||.:||||::|||:||||||   :.|.    ..   :::.
Qy  257  VDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALAD--MSSGK-Q--K-NQLV 309

Db  469  PYRESVLTWLLRENLGGNSKTVMLATISPASIHADETLATLRYACKARSIVNRVKVNESP 528
         |||:||||||| :.||||| |.|:| |||| |. |||.||||||.:.||.:.|..|||.|
Qy  310  PYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKNHAVVNEDP 369

Db  529  HDKIIRDLRAEVDRLKS-LRNEYERQRRL--SGNS-NNPVPRKIIIET---SVDETE--V 579
         . :::.:|: |::.|:| |:.:        .|..| .||.  .:.       ||:.   |
Qy  370  NARMIRELKEELAQLRSKLQSSGGGGAGGSGGPVEESYPPDTPLEKQIVSIQQPDATV 429

Db  580  EAL-RQQLAER--ERE-LSRA-QKSWMEKLKEAEDQRKSELRVLKRRGLALE--LTAE-Q 631
         : : :.::.::  .:. |.:. ||:..::| : .: :|. .| :.. |::.   :..: :
Qy  430  KKMSKAEIVEQLNQSEKLYRDLNQTWEEKLAKTEEIHKEREAALELGISIEKGFVGPYH 489

Db  632  -KQ-ACLVNLTADPILSGTLFYLLPQGLVRIGRGRLPGGSSSSQPDIVLDGPLVALQHCS 689
          ::  .|VNL:.|| ||  | | :.: .|: | . . .    . :.: :.|.|.:: :|:
Qy  490  SKEMPHLVNLSDDPLLAECLVYNIKPGQTRVG-N-V---NQDTQAEIRLNGSKILKEHCT 544

Db  690  IEHERGGKLYVIPGSEDFETYVNGELLKDRRQLFHGDRLVIGGSHYFRISNP 741
         :|:. |.|||||  :::|||||||: :|| .:::.|  |:: | .|:|| ||
Qy  545  FENV-DNVVTIVP-NEKAAVMVNGVRIDKPTRLRSGYRIILGDFHIFRFNHP 594
```

```
RESULT   15
ENTRY    T13750      #type complete
TITLE    kinesin-like protein 38B - fruit fly (Drosophila
                     melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE     13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change
         13-Aug-1999
ACCESSIONS T13750
REFERENCE  217752
#authors  Ohkura, H.; Torok, T.; Tick, G.; Hoheisel, J.; Kiss, I.;
          Glover, D.M.
#journal  J. Cell Sci. (1997) 110:945-954
#title    Mutation of a gene for a Drosophila kinesin like protein,
          KLP38B, leads to failure of cytokinesis.
#accession T13750
##status  preliminary; translated from GB/EMBL/DDBJ
##molecule_type mRNA
##residues 1-1121 ##label OHK
##cross-references EMBL:Y10667; NID:e1054562; PID:e1192009;
                   PIDN:CAA71675.1
GENETICS
#map_position 2
SUMMARY  #length 1121 #molecular-weight 125194 #checksum 7078

Query Match  15.8%;  Score 857;  DB 2;  Length 1121;
Best Local Similarity 37.3%;  Pred. No. 5.60e-125;
Matches 243;  Conservative 147;  Mismatches 179;  Indels 83;  Gaps 42;
```

```
Db  115  VSEESNMTVAVRVRPLNALECTRGQVTNVVQVHGNSNELTVQAGSSADAS-AG--VTH-- 169
         .|  .|: . |||||:||.  | .    : :.:.|| .::   ..||. |  ||   |:
Qy    1  MSGGGNIKVVVVRVRPFNAREIDRG-AKCIVRMEGNQTILTPPPGAEEKARKSGKTIMDGP 59
```

```
Search completed: Mon Aug 21 15:40:03 2000
Job time : 44 secs.
```

```
            ##residues      1-294 ##label COL1
            ##cross-references GB:U00996
        #accession   S72552
            ##molecule_type protein
            ##residues      247-264 ##label COL2
CLASSIFICATION  #superfamily kinesin-related protein KIF3; kinesin motor
                domain homology
KEYWORDS        P-loop
FEATURE
    1-279           #domain kinesin motor domain homology (fragment) #label
                    KMOT\
    29-36           #region nucleotide-binding motif A (P-loop)
SUMMARY         #length 294 #checksum 4908

Query Match         16.6%;    Score 900;  DB 2;  Length 294;
Best Local Similarity 48.7%;  Pred. No. 7.69e-133;
Matches  148;  Conservative   66;  Mismatches  77;  Indels  13;  Gaps  8;

Db     1 KQIDLYDETFRSLVESVLQGFNGTIFAYGQTGTGKTFTMEGVRSNPELRGVIPNSFEHIF 60
         :|:||:::  :|::::  :| |||||||||::|:::|  : |:||   :::|||  :
Qy    80 RQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-GK-E-HGVIPRICQDMF 136

Db    61 THIARTQSQQFL---VRASYLEIYQEEIRDLLAKDQKKRLDLKERPDTGVYVKDLSSFVT 117
         :|:  ||  |     ||||||||||||| | ||| || ||::|
Qy   137 RRINELQKDKNLTCVEVSYLEIYNERVRDLLNPSTKGNLKVREHPSTGPYVEDLAKLVV 196

Db   118 KSVKEIEHVMTVGNNNRSVGSTNMNEHSSRSHAIFPIITIECSELGVDGENHI-RYGKLNL 176
         ::|||:| :|:  ||:|:|:|| ::|||||||||| |:|
Qy   197 RSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHDEETKMDTEKVAKISL 256

Db   177 VDLAGSERQAKTGATGDRLKEATKINLSLSALGNVISALVD---GK--SSH-IPYRDSKL 230
         |||||||  |  ||| ||||| :| || ||||||||:
Qy   257 VDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSGKQKKNQLVPYRDSVL 316

Db   231 TRLLQDSLGGNAKTVMVANMGPASYNFDETITTLRYANRAKNIKNKPKINEDPKDALLRE 290
         | ||||| ||| |  |  :|||:||::|| ||||||:||||||:||||:
Qy   317 TWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKNHAVVNEDPNARMIRE 376

Db   291 FQEE 294
         : ||
Qy   377 LKEE 380


RESULT          13
ENTRY           A55236        #type complete
TITLE           kinesin-related protein KLP68D - fruit fly (Drosophila
                melanogaster)
ALTERNATE_NAMES kinesin-like protein 5; KLP5
ORGANISM        #formal_name Drosophila melanogaster
DATE            23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change
                22-Jun-1999
ACCESSIONS      A55236; E41298
REFERENCE       A55236
    #authors        Pesavento, P.A.; Stewart, R.J.; Goldstein, L.S.B.
    #journal        J. Cell Biol. (1994) 127:1041-1048
    #title          Characterization of the KLP68D kinesin-like protein in
                    Drosophila: possible roles in axonal transport.
    #cross-references MUID:95050960
    #accession      A55236
        ##molecule_type mRNA
        ##residues      1-784 ##label PES
        ##cross-references GB:U15974; NID:g595912; PIDN:AAA69929.1; PID:g565090
REFERENCE       E41298
    #authors        Stewart, R.J.; Pesavento, P.A.; Woerpel, D.N.; Goldstein,
                    L.S.B.
    #journal        Proc. Natl. Acad. Sci. U.S.A. (1991) 88:8470-8474
    #title          Identification and partial characterization of six members of
                    the kinesin superfamily in Drosophila.
    #cross-references MUID:92020874
    #accession      E41298
        ##molecule_type DNA
        ##residues      'TC',222-337,'VRQQV' #label STE
        ##cross-references GB:M74431; NID:g157791; PIDN:AAA28658.1; PID:g157792


GENETICS
    #gene           FlyBase:Klp68D; KLP5
    ##cross-references FlyBase:FBgn0004381
FUNCTION
    #description    may be part of a motor protein that provides anterograde fast
                    axonal transport
CLASSIFICATION  #superfamily kinesin-related protein KIF3; kinesin motor
                domain homology
KEYWORDS        ATP; coiled coil; microtubule binding; P-loop
FEATURE
    1-349           #domain head globular #status predicted #label HGL\
    20-350          #domain kinesin motor domain homology #label KMOT\
    106-113         #region nucleotide-binding motif A (P-loop)\
    350-580         #domain helical rod #status predicted #label ROD\
    581-784         #domain tail globular #status predicted #label TGL\
    112             #binding_site ATP (Lys) #status predicted
SUMMARY         #length 784 #molecular-weight 88193 #checksum 3313

Query Match         16.5%;    Score 897;  DB 1;  Length 784;
Best Local Similarity 46.5%;  Pred. No. 2.73e-132;
Matches  152;  Conservative   83;  Mismatches  72;  Indels  20;  Gaps  14;

Db    63 QRKVFTYDAAYDASATQTLYHEVVFPLVSSVLEGFNGCIFAYGQTGKTFTMEGVRGN 122
         ::|:||:    :|:: :::||: :|:::  |:: ||| ||||||||:: :::| |
Qy    65 DRSYWSFDKNAPNYARQEDLFQDLGVPLLDNAFKGYNNCIFAYGQTGSGKSYSMMGY-GK 123

Db   123 DELMGIIPRTFEQIWLHIN--RTE-NFQFLVDVSYLEIYMEELRDLLKPNSK-HLEVRER 178
         : |:|  |::|:|::|  :   ||| :|::|| |||||:: :::||
Qy   124 EH--GVIPRICQDMFRRINELQKDKNLTCVEVSYLEIYNERVRDLLNPSTKGNLKVREH 181

Db   179 -GSGVYVPNLHAINCKSVEDMIKVMQVGKNRTVGFTNMNEHSSRSHAIFMIKI-EM-CD 235
          ::| |:  ::  ::|: :|::::   :|||  ::|||:|||||||||:
Qy   182 PSTGPYVEDLAKLVVRSFQEIENLMDEGNKARTVAATNMNETSSRSHAVFTLTLTQKWHD 241

Db   236 TET--NTIKVGKLNLIDLAGSERQSKTGASAERLKEASKINLALSSLGNVISALAE-SS- 291
         || | |||:| :|:||||||| :|||:|| :|||:|::::|
Qy   242 EETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSTLGRVIAALADMSSG 301

Db   292 -P--H-VPYRDSKLTRLLQDSLGGNSKTIMIANIGPSNYNYNETLTTLRYGSRAKSIQN 346
          |  | |||||| ||:|||||||||| | |||::||:|:| :|| |||| :|:
Qy   302 KQKKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAAISPADINFEETLSTLRYADSAKRIKN 361

Db   347 QPIKNEDPQDAKLKEYQEEIERIK-RL 372
         :|:||| |::|||| ||:: |: ||
Qy   362 HAVVNEDPNARMIRELEKEELAQLRSKL 388


RESULT          14
ENTRY           A53939        #type complete
TITLE           kinesin homolog KHP1 - Chlamydomonas reinhardtii
ORGANISM        #formal_name Chlamydomonas reinhardtii
DATE            07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
                24-Sep-1999
ACCESSIONS      A53939
REFERENCE       A53939
    #authors        Walther, Z.; Vashishtha, M.; Hall, J.L.
    #journal        J. Cell Biol. (1994) 126:175-188
    #title          The Chlamydomonas FLA10 gene encodes a novel
                    kinesin-homologous protein.
    #cross-references MUID:94299638
    #accession      A53939
        ##status        preliminary
        ##molecule_type mRNA
        ##residues      1-786 ##label WAL
        ##cross-references EMBL:L33697; NID:g497696; PIDN:AAA21738.1;
                        PID:g497697
    ##note          authors translated the codon AAC for residue 753 as Asp
GENETICS
    #gene           FLA10
CLASSIFICATION  #superfamily unassigned kinesin-related proteins; kinesin
                motor domain homology
KEYWORDS        ATP; coiled coil; heptad repeat; P-loop
FEATURE
```